INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



PISTAS RAPIDAS

Febrero 1979

170 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 29
Febrero 1979



EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

José Manuel García de la Mora: *Matrimonio y fertilidad en los países desarrollados*; Pedro Pascual: *Moléculas nucleares*; José Costa: *Centrales energéticas de células de combustible*; Jesús Palomeque: *La estructura de la hemoglobina y el transporte respiratorio*; Monique Robert: *Ojos animales donde la imagen se forma mediante espejos*; Antonio Blanco: *El templo arrasado de Akhenaton*; Manuel Puigcerver: *Nebulosas gaseosas*; Luis Bou: *Juegos matemáticos*; José E. Myro y Alicia Menéndez: *Taller y laboratorio*.

Libros:

Lorenzo Alvarez, José Argemí y J. Antonio Campá



LA PORTADA

La portada representa el pie de un corredor al rebotar en la superficie de una nueva pista cubierta de atletismo, construida en la Universidad de Harvard. La infraestructura de la pista está hecha principalmente de madera y de un revestimiento superficial de poliuretano. Es mucho más elástica que las pistas normales. Por tanto, resulta más cómodo correr por ella y hacerlo más de prisa, batiendo los récords conseguidos hasta ahora (véase "Pistas rápidas", en *Ciencia y sociedad* de este mismo número). Se supone que las pistas blandas son las más lentas, y existe la creencia general de que las pistas duras son las más rápidas. Sin embargo, a través del análisis de un modelo de mecanismos de la carrera, se ha determinado que hay un punto de dureza intermedia en la que la velocidad del corredor aumenta por encima de su valor en superficie dura.

Suscripciones:
Prensa Científica, S.A.
Calabria 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 41

Condiciones de suscripción:
España:
Un año (12 números): 1870 pesetas
Extranjero:
Un año (12 números): 36 U.S. $
Ejemplar atrasado ordinario:
190 pesetas
Ejemplar atrasado extraordinario:
270 pesetas

Distribución para España:
Distribuciones de Enlace, S.A.
Ausias March, 49; Barcelona-10

Distribución para los restantes países:
Editorial Labor, S.A.
Calabria, 235-239; Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137; Madrid-30
Tel. 439 41 50 ext. 468
Cataluña:
Miguel Munill
Balmes, 191, 2.°, 3.ª; Barcelona-6
Tels. 218 44 45 y 218 40 86

Controlado por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Dibujo de la portada de Robert Giusti

Página	Fuente	Página	Fuente
7-12	Ilil Arbel	69-70	Bunji Tagawa
15	Ronald Y. Cusson Instituto de Física Teórica, Universidad de Frankfurt	71	Akhenaten Temple Project
16	Dan Todd	72-76	Bunji Tagawa
17	D. Allan Bromley, Yale University	77	Akhenaten Temple Project
18-19	Dan Todd	78	Bunji Tagawa
20	Yale University	81-82	Kitt Peak National Observatory
21-23	Dan Todd	83	Harvard College Observatory
24	Hubert C. Flocard, Steven Koonin y Morton Weiss	84	Albert Miller
29	Ben Rose	85	Haystack Observatory, Northeast Radio Observatory Corporation
30-34	Gabor Kiss	86	Albert Miller
41	Richard J.Feldmann y Thomas K. Porter, National Institutes of Health	87	Kitt Peak National Observatory
44-55	George V. Kelvin	88	Haystack Observatory, Northeast Radio Observatory Corporation
57	Michael F. Land, Universidad de Sussex	90	Harvard College Observatory
58	Tom Prentiss	91	Albert Miller
59	Vernon Barber, Memorial University of Newfoundland	92	Harvard College Observatory
60	Michael F. Land (*arriba, derecha*) y Tom Prentiss (*abajo, derecha*)	93	Albert Miller
61-65	Tom Prentiss	94	Albert Miller (*arriba*), Kitt Peak National Observatory (*abajo*)
66-67	David Scharf	99-105	E. Muñoz y J.M. Andreu
		109-110	Ilil Arbel
		112	Dover Publications, Inc.
		113-117	Michael Goodman

Dep. legal: B. 38.999-76
Fotocomposición: Tecfa
Guipúzcoa, 36 (local 1) Barcelona-20
Cayfosa. Santa Perpetua de Moguda
Barcelona.
Printed in Spain - Impreso en España

Los autores

CHARLES F. WESTOFF ("Matrimonio y fertilidad en los paises desarrollados") es catedrático del departamento de estudios demográficos y sociología de la Universidad de Princeton y director del seminario de investigación sobre población de dicho centro superior. Obtuvo su licenciatura en la Universidad de Siracusa y se doctoró por la de Pennsylvania en 1953. Ha estado en la facultad de Princeton desde 1955; desde 1970 a 1972 fue director ejecutivo de la Comisión sobre Crecimiento Poblacional y Futuro de los Estados Unidos.

D. ALLAN BROMLEY ("Moléculas nucleares") enseña física en la cátedra Henry Ford II de la Universidad de Yale, en donde dirige el laboratorio A. W. Wright de estructura nuclear. Nació en Ontario, se graduó en la Queen's University de Kingston en 1948 y se graduó en física el año siguiente. A continuación, se incorporó en el programa doctoral de la Universidad de Rochester, obteniendo su doctorado en 1952. Dos año después pasó al equipo de investigación de Atomic Energy of Canadá, Ltd., en donde colaboró en el proyecto del primer ciclotrón de energía variable. En 1960 se fue a Yale; allí, desde 1963 hasta 1965 participó en el diseño del primer acelerador electrostático Van de Graaff, de tipo tándem, modelo "Emperor" y supervisó su instalación en el Wright Laboratory.

ARNOLD P. FICKETT ("Centrales energéticas de células de combustible") es director de programación de sistemas de células de combustible y de energía química en el Electric Power Research Institute (EPRI) en Palo Alto, California. En 1956 se graduó en química en el Bates College, y entró a formar parte del equipo de investigación de la General Electric Company. A lo largo de sus 18 años en la General Electric desarrolló una amplia variedad de células de combustible. Ingresó en el EPRI en 1974 y desde entonces ha dirigido la investigación centrada en la tecnología de las células de combustible de carbonato fundido y de ácido fosfórico.

M. F. PERUTZ ("La estructura de la hemoglobina y el transporte respiratorio") es director del Consejo de Investigación Médica del Laboratorio de Biología Molecular de Cambridge, Inglaterra. Nació en Viena en 1914, e hizo la licenciatura en química en su Universidad. En 1936 se trasladó a Inglaterra para investigar bajo las órdenes de J. D. Bernal en el Laboratorio Cavendish de la Universidad de Cambridge. Después de obtener su doctorado en cristalografía de rayos X en 1940, trabajó como ayudante de investigación de W. L. Bragg en el Laboratorio Cavendish. Empezó su trabajo de la estructura de la hemoglobina en 1937, pero no fue hasta 15 años más tarde, en 1953, cuando finalmente descubrió un método para dilucidar la estructura de las moléculas proteicas. Su método condujo a la resolución de la estructura de la mioglobina por J. C. Kendrew y la de la hemoglobina por él mismo. Por estos descubrimientos Pertuz y Kendrew compartieron el premio Nobel de química en 1962.

MICHAEL F. LAND ("Ojos animales donde la imagen se forma mediante espejos") es actualmente profesor de neurobiología en la Universidad de Sussex. Nació en el Devon, realizó sus estudios de licenciatura en la rama de zoología en la Universidad de Cambridge, interesándose de forma especial por "la formas más extrañas de vida animal". Tras graduarse, trabajó en neurofisiología en el University College London, donde obtuvo el título de doctor en 1967 con una tesis sobre el ojo de la venera o concha de peregrino. Después de pasar dos años en calidad de becario postdoctoral en la Universidad de California en Berkeley, donde colaboró con Gerald Westheimer en el estudio de los movimientos de los ojos de las arañas, pasó a ocupar un puesto de profesor ayudante de fisiología y anatomía en Berkeley. Se trasladó a la Universidad de Sussex en 1971.

DONALD B. REDFORD ("El templo arrasado de Akhenaton") es profesor de estudios del Próximo Oriente en la Universidad de Toronto y asociado de investigación en el Museo Universitario de la Universidad de Pennsylvania. Se formó en Toronto, donde obtuvo la licenciatura en hebreo y acadio y su grado de doctor en estudios del Próximo Oriente en 1965 con una tesis sobre la cronología de la Dinastía XVIII egipcia. En la década de 1960 participó en las excavaciones de la Escuela Británica de Arqueología, en Jerusalén, bajo la dirección de Dame Kathleen Kenyon. De 1970 a 1972 fue director de una expedición al templo de Osiris en Karnak, y de 1972 a 1975 dirigió las excavaciones de Karnak-Este, patrocinadas por la Universidad de Pennsylvania.

ERIC J. CHAISSON ("Nebulosas gaseosas") es profesor agregado de Astronomía en la Universidad de Harvard y astrofísico del Centro de Astrofísica del Observatorio de Harvard College y del Observatorio Astrofísico Smithsoniano. Inicialmente especializado en física del estado sólido en las Universidades de Lowell y Harvard, obtuvo su doctorado en Astrofísica en 1972 en Harvard. Su trabajo de investigación se ha concentrado principalmente en el examen de los procesos físicos y químicos dentro de las nubes interestelares mediante la espectroscopía de radiofrecuencia. Recibió en 1977 el Premio Bart J. Bok por contribuciones sobresalientes al conocimiento de la galaxia, y fue investigador de la fundación Alfred P. Sloan desde 1976 a 1978. En la actualidad, Chaisson explica un curso sobre "Evolución cósmica" en el Harvard College, en el que se intenta sintetizar la astrofísica y la bioquímica.

EMILIO MUÑOZ y J. M. ANDREU ("Mecanismo molecular de la transducción de energía") trabajan en la Sección de Bioquímica de Membranas, adscrita al Instituto de Inmunología y Biología Microbiana del CSIC. E. Muñoz dirige dicha sección desde su creación en 1968. Estudió en las Universidades de Valencia y Madrid, graduándose en la última como doctor en farmacia en 1964. Trabajó durante cuatro años en el Departamento de Bacteriología de la Universidad de Lieja y durante dos años en el Departamento de Microbiología de la Universidad de Nueva York. En su investigación se interesa por el estudio de las paredes y membranas celulares en bacterias y de la membrana celular en eucariontes con especial énfasis en la caracterización de la dinámica molecular de los constituyentes de dichas estructuras en relación con sus funciones biológicas. J. M. Andreu es doctor en Ciencias Biológicas por la Universidad de Madrid en 1976. Desde 1971 ha trabajado en la Sección de Bioquímica de Membranas como becario predoctoral y postdoctoral (Beca Fundación March) y finalmente como contratado en la caracterización de ATPasas transductoras de energía. En 1976 se desplazó al Instituto Max Planck de Inmunología de Friburgo y actualmente se encuentra trabajando en el Departamento Graduado de Bioquímica de la Universidad Brandeis.

Hace...

Jose M.ª López Piñero

...cuatrocientos años

Juan de Herrera fue nombrado "aposentador del Real Palacio" de Felipe II, además de ser "architecto general de su Magestad". Desde ambos puestos, se convirtió en una de las personas de confianza del monarca en cuestiones técnicas y científicas.

Pesó decisivamente en la creación de la Academia de Matemáticas de Madrid a finales de 1582. La nueva institución, de la que fue el primer director, estuvo principalmente consagrada a la enseñanza de la cosmografía, la náutica, la arquitectura y la ingeniería.

Además de una obra arquitectónica sobradamente conocida, Herrera realizó una importante labor como ingeniero. Entre sus invenciones, figuran una serie de grúas de gran tamaño para el transporte y elevación de los enormes pesos que fue necesario mover en la construcción del Escorial. Ideó también un conjunto de ocho instrumentos "para hallar la longitud y latitud de las regiones en cualquier tiempo y hora del día y lugar donde se hallaren y ansí mesmo para averiguar lo que nordestea y noroestea la aguja tocada con la piedra imán". Cervera Vera ha dado a conocer recientemente que un "ingenio" suyo para cortar hierro alcanzó muy rentables resultados, instalado en una ferrería cerca de Durango.

Otro aspecto destacado de su actividad técnica fue la ingeniería hidráulica. Junto a otros muchos proyectos, participó en la planificación del pantano de Tibi, el más célebre de una serie de embalses valencianos y murcianos de los siglos XVI y XVII, cuyas dimensiones permanecieron sin igualar en el resto de Europa hasta la centuria ilustrada.

...trescientos años

Alfonso Limón Montero, catedrático de medicina en la Universidad de Alcalá, terminó su obra *Espejo Cristalino de las Aguas de España,* que no fue impresa hasta 1697, varios años después de su muerte. Se trata del gran título inicial de la hidrología médica en España, de especial interés por su contenido químico.

Para realizar su estudio, Limón Montero recurrió a casi medio centenar de corresponsales de toda la Península, que le proporcionaron datos acerca de las aguas mineromedicinales existentes en las localidades en las que residían. Casi todos eran médicos, algunos de los cuales practicaron destilaciones. A dicha información, añadió "lo que yo pude averiguar por mi persona". Pesó exactamente el residuo de las aguas para conocer su mineralización, examinó los caracteres organolépticos y las sometió a análisis químico.

Incluyó en su obra numerosos datos químicos, no solamente relativos a las aguas, sino también a la preparación de medicamentos y a las propiedades de diversos minerales y metales. Citó a Paracelso y a varios paracelsistas, a iatroquímicos y a otros autores relacionados con la química, como Agrícola. Se opuso abiertamente a las especulaciones alquímicas con argumentos experimentales y criticó agriamente "cuantas falsedades se experimentan en las medicinas así simples como compuestas, sólo por un corto interés o por ahorrarse algún trabajo en sus preparaciones".

...doscientos años

Apareció en Amsterdam el libro *Synopsis stirpium indigenarum Aragoniae* de Ignacio Jordán de Asso. El interés de este célebre jurista e historiador por las ciencias de la naturaleza resultó estimulado cuando se hizo cargo del consulado español en la citada ciudad holandesa (1776) y entabló relación directa con los naturalistas allí residentes. En la *Synopsis*, expuso los resultados de sus excursiones botánicas por el sur de Aragón, donde recogió más de un millar de especies botánicas. Más tarde, viajó por el monte Guara y la zona de Jaca, reuniendo materiales que resumió en una especie de apéndice de la obra anterior (1781). En un tercer viaje que efectuó por el Pirineo aragonés, se ocupó principalmente de los minerales de la zona, así como de sus fósiles o "petrificata". Su nuevo libro (1784) se recuerda sobre todo por su contribución a la paleontología, en un período durante el cual esta disciplina tuvo varios destacados cultivadores españoles.

Asso desarrolló, por otra parte, una intensa actividad en la Sociedad Económica de Amigos del País de Zaragoza. Regaló a esta institución sus colecciones de naturalista y formó parte de la comisión encargada de formar un jardín botánico. Dentro del marco de la Sociedad, publicó un estudio sobre la langosta y la forma de combatirla (1785) y realizó experiencias acerca del cultivo y el rendimiento de diversas clases de trigo (1785).

...cien años

Se realizó la unión geodésica y astronómica de los continentes europeo y africano bajo la dirección de Carlos Ibáñez e Ibáñez de Ibero.

Ingeniero militar de carrera, Ibáñez era entonces una de las grandes autoridades internacionales en la materia. Había dirigido la medición de la base central de la triangulación geodésica de España, en la provincia de Toledo, con una precisión que mereció que la Academia de Ciencias de París la calificara como "memorable operación científica que no podía ser superada" (1863). Ello fue posible, en gran parte, gracias a la utilización de un instrumento para medir bases geodésicas diseñado por él: la llamada "regla española". Este aparato acabó imponiéndose, hasta el punto de que sirvió de modelo al que la Asociación Geodésica Internacional mandó construir para medir las bases de las redes geodésicas de las distintas naciones europeas. Más tarde, su autor lo perfeccionó.

Ibáñez fue el primer director del Instituto Geográfico y Estadístico, creado por iniciativa suya en 1870 y cuya principal aportación inicial consistió en emprender la publicación de un mapa topográfico de España en escala de 1:50.000. Bajo su presidencia, se reunió en 1875 la conferencia de la que surgieron la Oficina y el Comité Internacional de Pesas y Medidas. También fue nombrado presidente de este último. Participó en otras importantes tareas científicas nacionales e internacionales y varias de sus publicaciones, entre ellas la titulada *Base central de la triangulación geodésica en España* (1865), fueron traducidas a numerosos idiomas.

...cincuenta años

Eduardo Vitoria Miralles fundó *Afinidad*, que hoy continúa siendo una de las principales revistas españolas consagradas a la química.

Vitoria terminó su licenciatura en ciencias físicas y químicas (1895), varios años después de su ingreso en la Compañía de Jesús. Completó su formación científica en la Universidad de Lovaina –donde obtuvo el doctorado– y en varias prestigiosas instituciones alemanas. De regreso a España, trabajó algún tiempo en Tortosa hasta que, en 1916, fue trasladado al Colegio de San Ignacio en Sarriá (Barcelona), con el encargo de que lo convirtiera en un centro de estudios químicos. Aquel mismo año fue inaugurado el Instituto Químico de Sarriá, que Vitoria organizó de acuerdo con el más exigente nivel europeo de la época. La revista *Afinidad*, antes citada, fue en su origen (1921) el boletín de la Asociación de Químicos del Instituto.

Matrimonio y fertilidad en los países desarrollados

El descenso del índice de natalidad en los países industrializados es causa y consecuencia de cambios en el matrimonio y en la familia. ¿Cómo afrontarán las naciones una disminución prolongada de la población?

Charles F. Westoff

La mayoría de los países desarrollados están rondando actualmente el cero de aumento de población. En Europa son varios –Austria, las dos Alemanias y Luxemburgo– los que andan ya por debajo de ese cero: tienen al año más muertes que nacimientos. En el Reino Unido la balanza está, de momento, equilibrada, con un número casi igual de nacimientos y defunciones. Si siguen dándose las actuales tendencias de la fertilidad, Bélgica, Checoslovaquia, Dinamarca, Hungría, Noruega y Suecia llegarán en pocos años al cero de crecimiento o caerán por debajo de él; y hacia 1990 las habrán seguido en este descenso Bulgaria, Finlandia, Grecia, Italia y Suiza. De persistir las mismas tendencias, para el cambio de siglo las poblaciones de Europa y de la Unión Soviética empezarían en conjunto a disminuir y, hacia el año 2015, la de los Estados Unidos dejaría de aumentar.

¿Continuará disminuyendo la fertilidad? Acerca de esto sólo caben especulaciones más o menos plausibles. Aquí voy a sostener que el bajo índice de fertilidad es el resultado de una tendencia a largo plazo que tal vez se siga dando. Y examinaré algunos de los cambios sociales que se están produciendo actualmente –en particular los que afectan a la institución del matrimonio, que parecen la causa probable de los recientes descensos de la fertilidad en los países industrializados– y algunas de las maneras como pueden estos países afrontar una disminución real o inminente de la población.

El pronóstico según el cual alrededor del año 2000 mermará la población europea se basa en el supuesto de que la tasa total de fertilidad en Europa llegará, hacia 1986, al 1,5 y se mantendrá a ese nivel. (La tasa total de fertilidad en un año determinado es el número promedio de hijos que tendrían las mujeres a lo largo de toda su vida reproductiva si se ajustasen a las tasas características de los diversos grupos de edades para ese año; se calcula sumando las tasas de nacimiento características para las diversas edades en un año dado. La tasa total de fertilidad con la que una población se renueva y autorremplaza es 2,1 más bien que 2,0; en primer lugar, porque no todas las hembras sobreviven hasta alcanzar las edades reproductivas. La proporción relativamente grande de población juvenil con que cuentan hoy muchos países significa que han de transcurrir algunos años de fertilidad inferior a la necesaria para el adecuado remplazo, antes de que una población deje de aumentar y empiece a disminuir.) Una tasa total de fertilidad del 1,5 es inferior a las tasas predominantes hoy día en Europa, y queda por debajo de la menor de cuantas se sabe que se hayan dado realmente en las mujeres de un amplio grupo generacional (1,8 para las mujeres nacidas en 1907 en Inglaterra y en Gales). Sin embargo, la tendencia de la fertilidad está siguiendo hoy una trayectoria que para 1986 la reduciría al 1,5. La mayor disponibilidad de complejas técnicas de control de natalidad hace cada vez más posible que se llegue a tan bajo nivel de fertilidad; todo parece apuntar en esa dirección. (Puede que la trayectoria de los futuros índices de reproductividad tienda a descender demasiado y con excesiva rapidez, en cuyo caso la perspectiva del tiempo quizá quede rebasada en pocas décadas.)

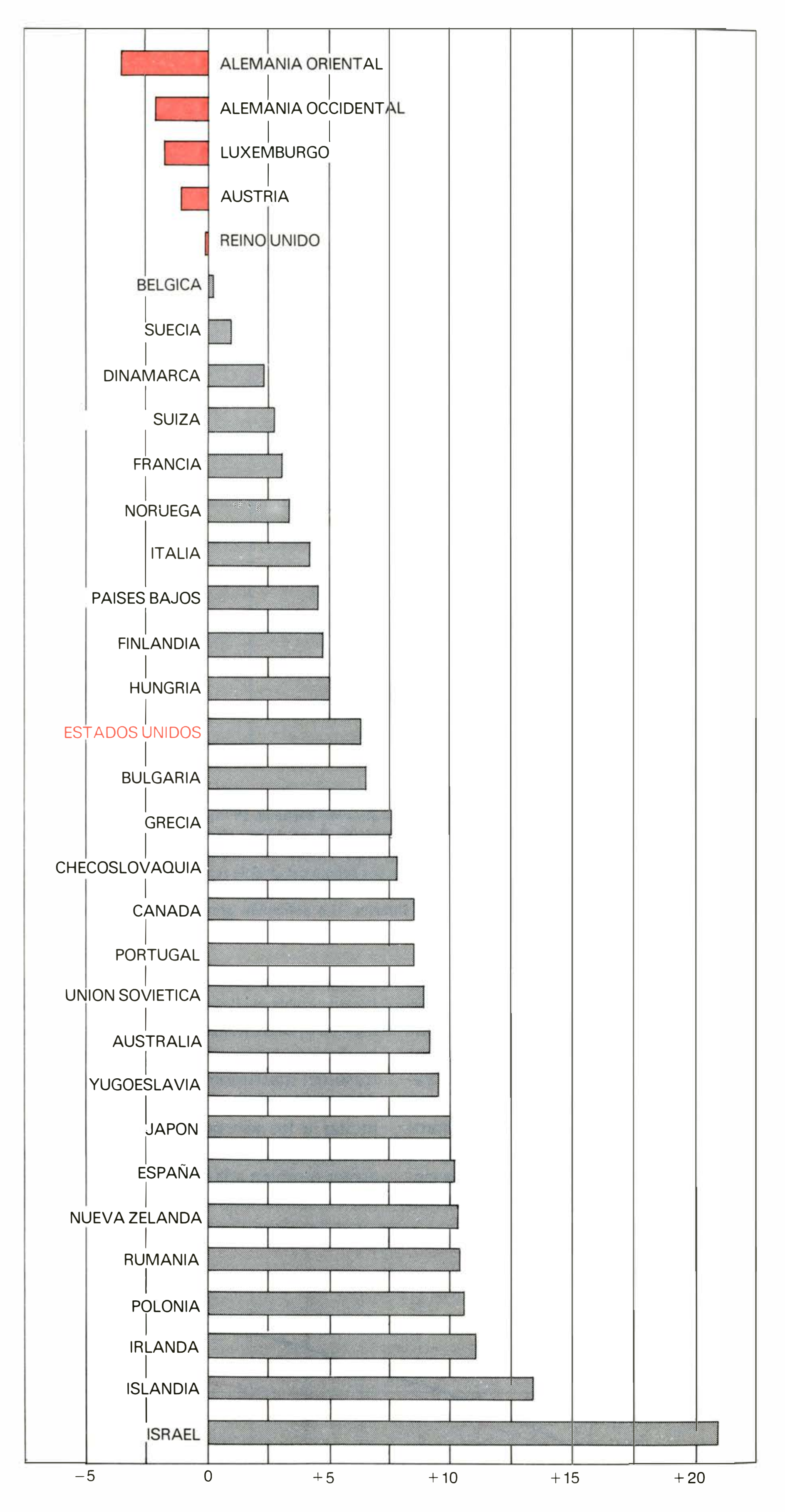
A causa de la sostenida explosión demográfica que siguió a la Segunda Guerra Mundial y del resultante incremento de la población apta para tener hijos, se puede esperar que los Estados Unidos (junto con Canadá, Australia y Nueva Zelanda) tardarán algo más que la mayoría de las naciones europeas en descender hasta el cero de aumento de población. Incluida la inmigración legal y suponiendo que el nivel básico de fertilidad sea, para un grupo generacional completo, de 1,7 partos por cada mujer (en los Estados Unidos es hoy corriente una tasa de reproductividad que implica alrededor de 1,8 nacimientos), la población norteamericana dejaría en efecto de aumentar cuando llegase a ser, en el año 2015, de unos 253 millones. Este es el pronóstico actual "bajo" de la Oficina del Censo. Aun suponiendo que la fertilidad volviera a subir hasta el nivel de remplazo del 2,1, la población no pasaría de los 260 millones en el año 2000, y sería de 283 millones en el 2015; es decir, bastante inferior a los 300 millones que hace solamente unos pocos años se predijeron para el 2000. Pero la diferencia importante entre el pronóstico basado en una tasa de fertilidad del 1,7 y el basado en otra del 2,1 no consiste tanto en esos 30 millones de diferencia para la población del año 2015 cuanto en que, con la tasa de fertilidad más alta, la población todavía seguiría aumentando cada año aproximadamente un 5 por 1000 del total. Ambos pronósticos prevén un considerable crecimiento más allá de la población actual –que anda un poco por debajo de los 220 millones– habida cuenta de la proporción relativamente alta de habitantes que, como consecuencia de la explosión demográfica, se hallan en edades aptas para tener hijos.

Una vez presentados estos pronósticos recientes, he de admitir que el registro de pronósticos poblacionales no ha sido muy afortunado. Así, el de 1947 sobre la población que tendrían los Estados Unidos en 1970 previó 65 millones menos de los 205 que en realidad se censaron ese año. Diferentes pronósticos hechos desde comienzos de la década de los sesenta, sobre la población de los Estados Unidos en el año 2000, han variado desde un máximo de 362 millones (así uno que se hizo en 1964) hasta un mínimo de 245 millones (la última estimación más baja). La principal variable determinante en tales cálculos es la fertilidad, y ésta ha cambiado mucho

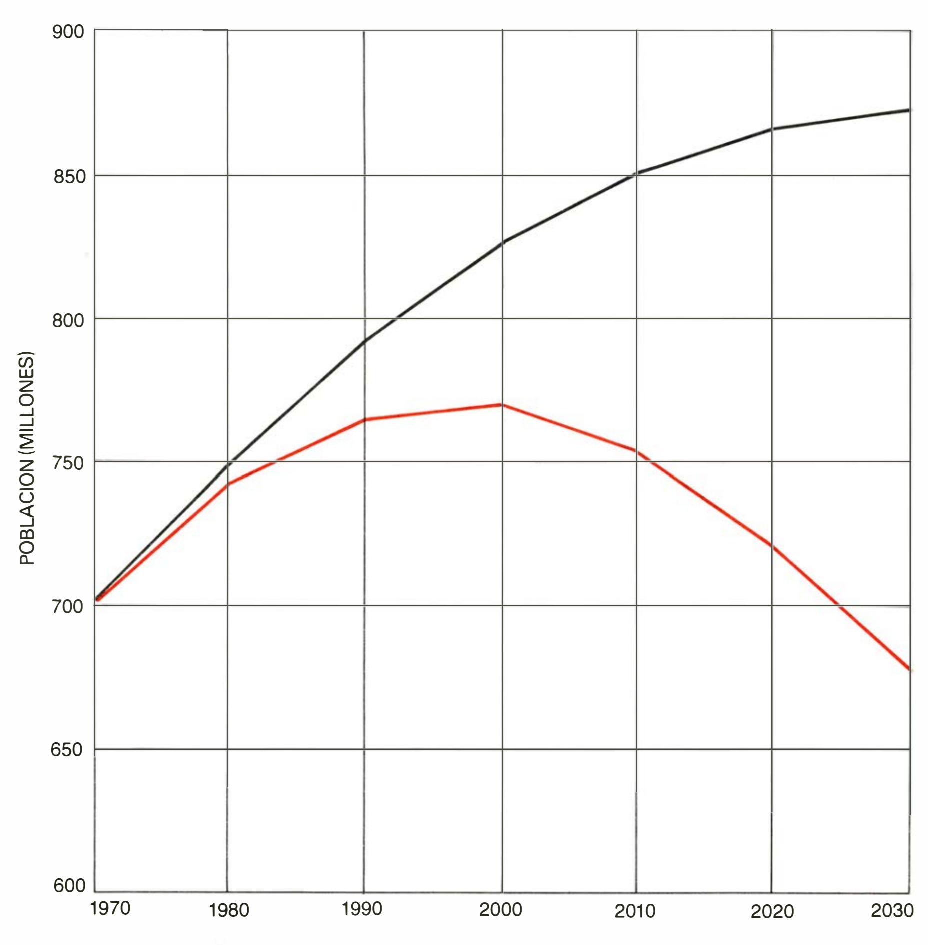
desde comienzos de la década de los 60. Si se advierte que todo pronóstico poblacional se basa esencialmente en una extrapolación de los niveles de fertilidad actuales, rodeados de una banda por "arriba" y otra por "abajo", no hay por qué extrañarse de que, durante un período de rápidos cambios del comportamiento reproductivo, los pronósticos no resulten muy certeros. Para poder predecir con más precisión, los demógrafos tendrían que saber mucho más que lo que saben sobre los determinantes sociales y económicos de la fertilidad. Y aun suponiendo tanta comprensión teórica, habría que prever además el curso que seguirían en el futuro los pertinentes índices sociales y económicos; ahora bien, no consta que las más relevantes disciplinas científico-sociales hayan desarrollado una tal capacidad.

Todo esto sugiere que es vana precipitación hasta el especular sobre las futuras tendencias de la fertilidad y del aumento de la población. ¿Cómo saber, en efecto, si se prepara o no una nueva explosión demográfica? De hecho, hay una escuela de pensamiento demográfico que ofrece una atrayente teoría en apoyo de una predicción justo así para los Estados Unidos. Su argumento es que la fertilidad responde positivamente a la percepción que la gente tiene de su grado de oportunidad, sobre todo si a la oportunidad le afecta la competición entre iguales. Tal competición es, en parte, cuestión de número. Quienes nacieron durante el período de bajo índice de nacimientos de los últimos años de la década de 1920 y toda la década de los 30, llegaron a su mayoría de edad después de la Segunda Guerra Mundial, en un período de economía en expansión con mucha demanda de fuerza laboral. Ellos fueron los padres de la explosión demográfica. En cambio, quienes nacieron en aquella explosión demográfica alcanzaron su mayoría de edad a finales de los años 60 y durante los 70, cuando su gran número hizo aumentar la competencia por los puestos escolares y laborales. Ellos, a su vez, han sido los causantes del retraso de los casamientos y del bajón de la fertilidad que han caracterizado la pasada década. Esta negativa asociación de la cantidad numérica de una generación y su fertilidad básica induce a predecir que se producirá otra gran explosión demográfica por los años 90, cuando las generaciones de los nacidos alrededor de 1970, período en el que la tasa de nacimientos fue baja, respondan a las más risueñas esperanzas y ventajas competitivas que les depare el hecho mismo de no ser tantos en número.

La teoría es, sin duda, muy sugerente; pero la evidencia empírica en pro de esos ciclos que predice redúcese casi a los dos ejemplos que acabamos de citar, y en



TASA DEL CRECIMIENTO NATURAL (tasa de nacimientos menos tasa de defunciones). Dicha tasa de crecimiento se halla por debajo de la línea del cero de aumento de población en unos pocos países europeos y próxima a ese nivel en la mayoría de los países industrializados. Los datos corresponden casi todos a 1976; para los Estados Unidos se tomaron los de 12 meses que concluyeron con el de mayo de 1978.



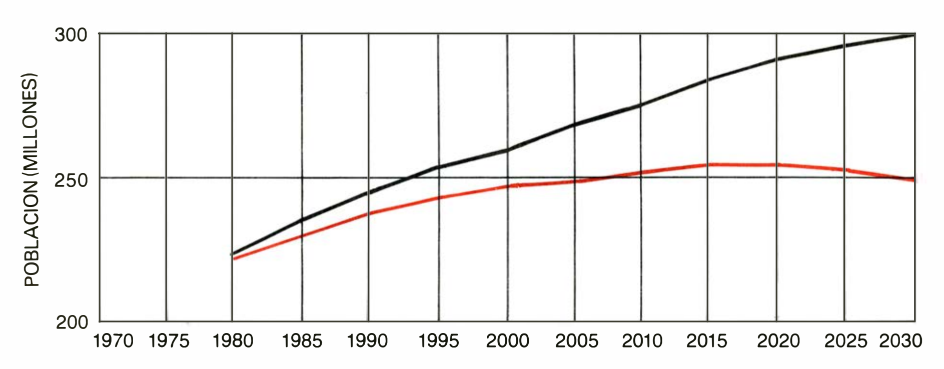
LA POBLACION DISMINUIRIA en Europa (incluyendo a la Unión Soviética) después del cambio de siglo, según un pronóstico de las Naciones Unidas *(trazo de color)* hecho en 1975; este pronóstico supone que la tasa total de fertilidad habría caído a 1,5, o sea, alrededor de un 25 por ciento por debajo de la tasa de remplazamiento, hacia el año 1986, y que se mantendría a ese nivel. Otro pronóstico de las Naciones Unidas *(trazo negro)* supone una fertilidad constante justo por debajo del nivel de remplazamiento, en cuyo caso la población seguiría aumentando, en proporción cada vez menor, durante el período representado. Las tendencias actuales autorizan más el pronóstico inferior. (La población prevista se cifra en millones.)

otros países tampoco se han podido aducir en su favor muchas más pruebas. Aparte de que tal teoría ignora los radicales cambios que están teniendo lugar en cuanto al status y a las expectativas de la mujer. Parece dar por descontado que la ventaja del número menor en la economía competitiva sólo habrán de hacerla realidad los varones, siendo así que las mujeres se vienen incorporando a la fuerza laboral en proporciones que cada día aumentan más de prisa. Ni es del todo claro el que una mejor oportunidad económica haya de traducirse automáticamente en menor retraso del matrimonio y en más alto nivel de fertilidad: ¿no podría, por el contrario, incitar a los jóvenes a gastar más en el ajuar y en superfluidades consumistas, y atraer también a las mujeres a prepararse para mejores empleos mediante el estudio de carreras largas? La teoría cíclica es, empero, coherente con alguna apoyatura empírica. ¿Qué pruebas abonan la opuesta opinión de que la fertilidad seguirá baja?

Cada vez que el índice de nacimientos desciende (y en estos últimos años ha descendido a menudo) los periodistas le preguntan al demógrafo a qué puede atribuirse tal descenso: ¿a "la píldora", al aborto, a la esterilización, a la recesión económica, al movimiento feminista o a alguna otra explicación ad hoc? Pero inquirir la causa del descenso más reciente equivale a no saber preguntar. Porque la caída del índice de natalidad es una realidad a largo plazo. Tal descenso se viene prolongando desde hace ya mucho: durante los últimos 200 años no ha dejado de bajar este índice más o menos de prisa en los Estados Unidos, con la excepción de un solo período. La pregunta certera, y la más intrigante, es la de cuál fue la causa de tal excepción, o a qué se debió la explosión demográfica que duró más de una década después de la Segunda Guerra Mundial.

Aquella explosión de nacimientos estamos empezando a comprenderla demográficamente, aunque ignoremos aún sus causas profundas. La noción de que todos se casaban entonces pronto y tenían muchos hijos es, cuando menos, imprecisa. Lo que, al parecer, produjo el estallido fue un movimiento muy marcado de huida de la soltería y de menor aspiración al matrimonio sin hijos o con un solo hijo. Aumentó mucho la proporción de mujeres que tuvieron por lo menos dos hijos, pero el proporcional aumento del número de las que tuvieron tres o más no desempeñó un papel tan importante en la explosión; no volvieron a darse las familias numerosas de otros tiempos. Qué fuerzas sociales y económicas generaran y sostuvieran los cambios en el matrimonio y en la procreación y crianza de los hijos no está del todo claro. Parece tener algo que ver con tales cambios un elemento imprevisible de lo que podría llamarse moda, lo cual es una razón para decir que tal vez pudieran darse de nuevo algún día los niveles de matrimonio y fertilidad que produjeron la explosión demográfica, por improbable que hoy parezca esto.

Si la explosión de nacimientos fue un excurso temporal (pero demográficamente significativo), la secular tendencia hacia abajo en cuanto a fertilidad, de la que se apartó, era un aspecto de una más amplia transición demográfica en Europa y los Estados Unidos: la de la disminución a largo plazo de los índices de mortalidad y de nacimientos que acompañó al desarrollo económico y la modernización en general. Lo básico para toda la transición demográfica ha sido un cambio en el valor

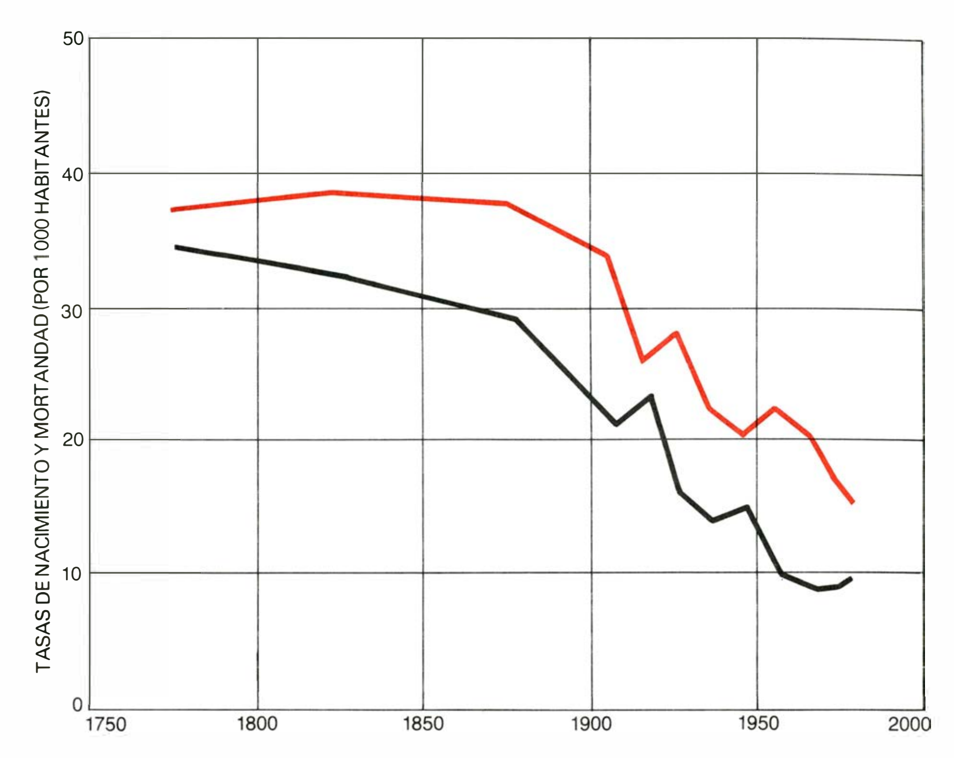


EN LOS ESTADOS UNIDOS el descenso de la población empezaría algo después, según un pronóstico de la Oficina del Censo *(trazo de color)* que supone una caída del total de la fertilidad al 1,7, por debajo de la actual tasa norteamericana, que es 1,8. Otro pronóstico *(trazo negro)* supone un aumento de la fertilidad hasta el nivel de remplazo adecuado, o sea, hasta el 2,1. Ambos pronósticos suponen una inmigración legal neta de 400.000 personas por año. Se desconoce la magnitud y la duración de la inmigración ilegal.

económico de los niños, que han pasado de ser una fuente de ingresos en las economías agrarias a ser, en las industriales y postindustriales, una considerable sangría de la renta. La transformación económica de la sociedad ha venido acompañada de un declinar de la autoridad tradicional y religiosa, y ha traído también la difusión de un ethos de racionalidad e individualismo, la educación a todos los niveles de ambos sexos, la creciente igualdad de la mujer, la supervivencia cada vez más numerosa de los hijos y la pujanza de una civilización consumista que cada día se orienta más a maximalizar la gratificación personal. Cuando estos cambios se combinan con el desarrollo y la difusión de una compleja tecnología de control de natalidad, ¿cómo extrañarse de que las instituciones del matrimonio y la familia muestren asimismo señales de cambio?

Lo importante en torno a este catálogo de cambios sociales es que todos los cambios parecen ser irreversibles y que algunos de ellos, particularmente los que atañen al status de la mujer, todavía no se han efectuado por entero. A pesar de ciertos indicios manifiestos de migración hacia las zonas rurales, no es probable un retorno a la economía agraria. Se está dando un cierto resurgir de la religión evangélica, pero nadie asegurará por eso que se vaya a restablecer la autoridad religiosa o tradicional de pasados siglos. Pese a algunos recientes desencantos respecto a las retribuciones económicas de la educación superior, se ha hecho normal el prolongar los años de educación y aprendizaje. Al economista profesional no le afecta de hecho el saber que el ansia de bienes de consumo no ha disminuido a pesar de los sentimientos antimaterialistas del movimiento en pro del medio ambiente. La tecnología del control de nacimientos, aunque todavía dista mucho de ser la ideal, se emplea hoy más amplia y eficazmente que en ninguna época anterior. Los actuales esfuerzos por invertir la tendencia a que se permita el aborto quizá tengan éxito en parte, pero puede que el resultado final se reduzca a hacer el aborto menos conveniente y más costoso. Nos estamos aproximando rápidamente a la "perfecta" sociedad anticonceptiva, en la que ya no habrá nacimientos no deseados.

Para abreviar, nada asoma por el horizonte que nos sugiera que la fertilidad no vaya a seguir a un bajo nivel. Todos los datos recientes sobre las tendencias del matrimonio y del comportamiento reproductivo inclinan a suponer que continuará siendo baja. Resulta instructivo considerar algunas de estas tendencias en los Estados Unidos y en dos países que han estado a menudo en cabeza del cambio social entre los países desarrollados: Dinamarca y Suecia.



MENGUA RECIENTE de la tasa de aumento de población en los países actualmente desarrollados, que resulta de un declinar de la tasa de nacimientos *(trazo de color)* mayor que el de la tasa de defunciones *(trazo negro)*. La actual convergencia de las tasas de nacimientos y de defunciones a un bajo nivel de fertilidad y de mortalidad indica que se está volviendo al equilibrio demográfico. Tal equilibrio había sido históricamente lo normal (aunque a un alto nivel de fertilidad y de mortalidad) antes de la segunda mitad del siglo XVIII.

El número de matrimonios ha ido disminuyendo en Dinamarca y Suecia a lo largo de la última década –el 30 por ciento en Suecia entre los años 1966 y 1975–. En los Estados Unidos ha habido una radical merma en la proporción de mujeres que se casan jóvenes. En 1960 sólo alrededor del 29 por ciento de las mujeres de 20 a 24 años de edad permanecían solteras; para 1978 la proporción se había elevado a casi el 48 por ciento. Esta proporción es aproximadamente la misma que en 1940, y es menor que a comienzos de siglo, cuando más del 50 por ciento de las mujeres de 20 a 24 años estaban aún solteras. Ulteriores aumentos de la soltería y de la posposición del matrimonio no dejarían de cuadrar –para decirlo de otro modo– con la experiencia histórica, pero el desarrollo contemporáneo sugiere que tales aumentos se producirán en el futuro por razones muy diferentes. Muchas mujeres de hoy retrasan el casarse hasta que frisan los 30. Algunos de estos retrasos se deben al temor a "la trampa del matrimonio": la rápida subida de la tasa de nacimientos después de la Guerra europea significó que a finales de los 60 y comienzos de los 70 había más mujeres dispuestas a casarse que hombres en la edad convencionalmente idónea ara hacerlo (unos pocos años mayores que ellas). Sin embargo, la posposición del matrimonio se debe en gran parte a cambios básicos en la actitud respecto al mismo en edades tempranas.

Abona esta conclusión la evidente y creciente tendencia a vivir emparejados sin casarse, sin formalizar el matrimonio. En los Estados Unidos se calculó en 1976 que un millón de parejas (alrededor del 2 por ciento de quienes vivían emparejados) no eran matrimonios. Parece cosa cierta que en el futuro próximo la cohabitación sin matrimonio se irá haciendo cada vez más popular. Su incidencia ha ido en rápido aumento en Dinamarca, donde más de la cuarta parte de las jóvenes de 18 a 25 años conviven con hombres con los que no están casadas. En Suecia un examen reciente hallaba que, de todas las parejas (de los 16 a los 70 años de edad) que hacen su vida en común, alrededor del 12 por ciento no están casadas. En muchos casos la cohabitación es simplemente un preludio para el matrimonio –una nueva forma de noviazgo–, y así la realidad quizá no haya cambiado mucho, aparte del legitimar la situación mediante un certificado. Por otra parte, puede que estemos ante el comienzo de un cambio más básico, que llegue a alterar significativamente la institución del matrimonio. En cualquiera de estos supuestos, lo que está claro es que la cohabitación trae consigo una mengua de la fertilidad, ya sea a causa del carácter de prueba propio de esta situación, ya por su relativa inestabilidad.

La inestabilidad de la relación matrimonial tiende cada vez más a ser la expectativa normal, a juzgar por el elevado y

creciente índice de divorcios, en los Estados Unidos. La tradicional noción de "compañeros para toda la vida" no es ya la corriente en un sector, cada vez más amplio, de la población norteamericana. El porcentaje actual de divorcios implica que quizá la mitad del total de matrimonios existentes llegarán algún día a divorciarse y que un tercio del total de niños tendrán que vivir bastante tiempo con uno de sus progenitores separado o divorciado. El porcentaje de divorcios ha venido a doblarse durante la pasada década en Dinamarca y en Suecia, aunque es todavía más bajo en estos países que en los Estados Unidos.

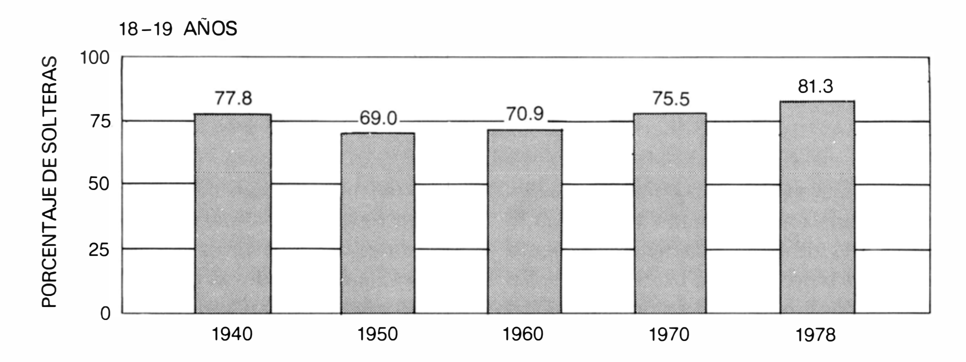
La opinión convencional sobre la elevada tasa estadounidense de divorcios es que expresa un desengaño no respecto al matrimonio en cuanto tal, sino respecto a un cónyuge determinado; adúcese en prueba de ello la alta proporción de re-casados. Mas también aquí puede discernirse el comienzo de un cambio: la proporción de re-casados, que en los Estados Unidos había ido aumentando a la par que la de divorciados, ha empezado a disminuir. (En Suecia la tasa de re-casamientos ha bajado casi un 50 por ciento desde 1965.) Muchos segundos matrimonios parecen ir hoy al divorcio. La cohabitación no está limitada a los años premaritales: cada vez se sabe de más divorciados de mediana edad e incluso mayores que cohabitan sin estar casados.

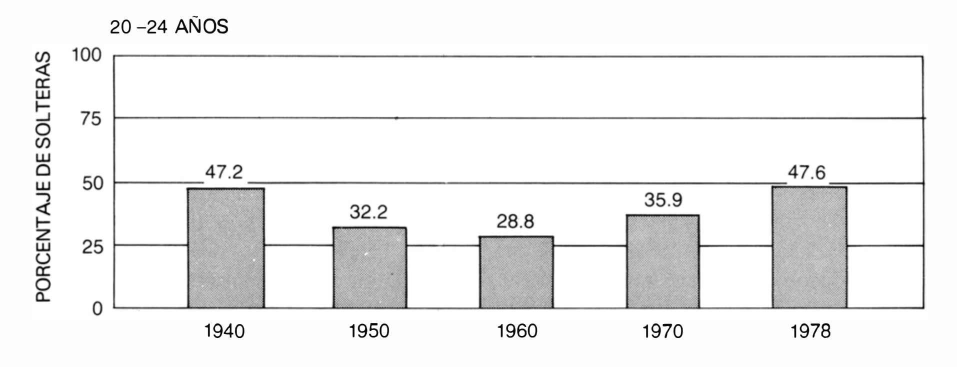
Entre los factores que han influido en el descenso de la fertilidad uno de los principales ha sido el cambio del papel de la mujer en la economía. Las últimas décadas han visto un pasmoso aumento de la proporción de mujeres incorporadas a la fuerza de trabajo. En los Estados Unidos la proporción de mujeres que durante los primeros años de la crianza de sus hijos trabajaban fuera de sus casas aumentó desde menos de 2/5 en 1960 a alrededor de 3/5 en 1977, y se cree que en otra década llegarán a ser 2/3. Para el conjunto de Europa, la Organización Internacional del Trabajo ha pronosticado que, hacia el año 2000, formarán parte de la fuerza laboral 3/5 del total de las mujeres que estén en edad reproductora.

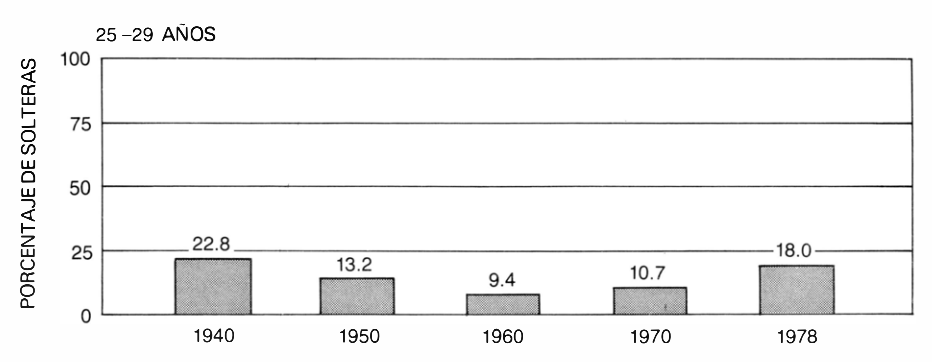
Tal aumento de la actividad económica por las mujeres no significa, a buen seguro, un aumento equivalente de su independencia financiera; sigue habiendo una clara desigualdad rentística entre los dos sexos. Las ganancias de las mujeres son bastante inferiores a las de los hombres, por más que se aminoren las diferencias educacionales entre los sexos. Las mujeres están desproporcionadamente concentradas en tareas menos remunerativas y aunque realicen idénticos trabajos que los hombres se las paga menos. Entre la población que trabajó a sueldo y por jornada completa durante todo el año, de edades de 18 a 44 años, en los Estados Unidos en 1976, la mujer promedio ganó solamente el 60 por ciento de lo que ganó el hombre promedio. Con todo, el sesgo futuro del status económico de la mujer parece claro: aunque la auténtica igualdad económica aún tardará probablemente varias generaciones en alcanzarla, cada día irá siendo mayor el número de mujeres que puedan optar por la independencia financiera.

Abundan los estudios dedicados a investigar las complejas interconexiones causales entre la participación de las mujeres en la fuerza de trabajo y la fertilidad de las mismas. El empleo y los hijos no siempre son incompatibles, y diversas secuencias de trabajo y de crianza complican la relación. Pero, sin meternos en calificaciones, parece bastante indudable que el trabajo y la fertilidad se correlacionan negativamente y que en el futuro aumentará la proporción de las mujeres que trabajan. Imaginemos, por último, una sociedad en la que hombres y mujeres tuvieran los mismos ingresos, en la que hubiese tantas mujeres como hombres ejerciendo de abogados, ingenieros, jueces, médicos, ejecutivos empresariales y en todos los demás cargos y profesiones: ¿cuáles serían las consecuencias para el matrimonio y la fertilidad?

Ciertamente la institución del matrimonio perdería una de las pocas razones de ser sociológicas que le quedan si se concibe el matrimonio, en términos funcionalmente económicos, como un sistema en el que la mujer ofrece sus servicios domésticos y de ama de cría a cambio de la seguridad y la situación que le proporcionen el trabajo y el sueldo o las rentas del hombre. Este concepto del matrimonio nada tiene de romántico, pero no hay duda de que contribuye mucho a explicar la universalidad y la persistencia de la institución. Si se eliminara la razón de ser económica, ¿qué motivaría entonces el matrimonio? ¡Hoy apenas es ya necesario





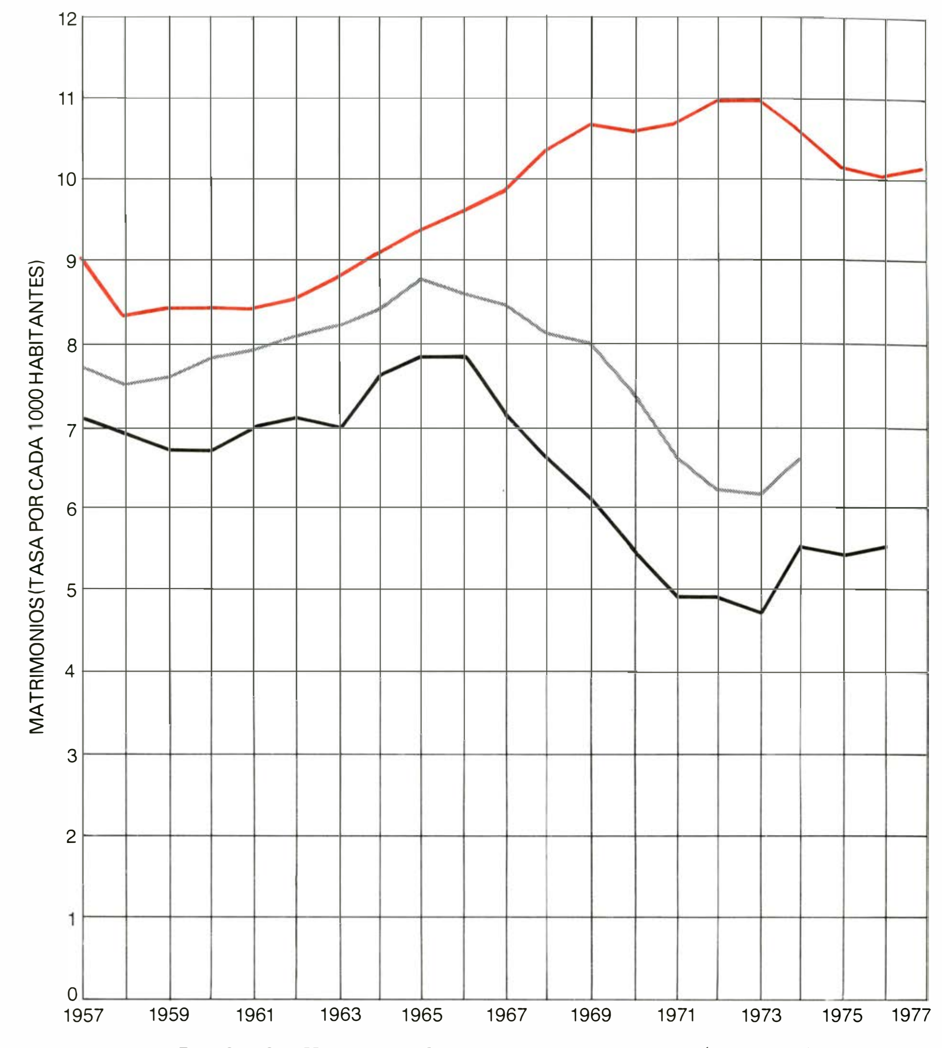


LA EDAD A QUE SE CASA LA GENTE afecta a la tasa de nacimientos. La explosión demográfica de los años 50 y primeros 60 llevó la marca del matrimonio temprano, según puede verse en estos rectángulos que indican los porcentajes de las mujeres que permanecían aún solteras a las diversas edades. El porcentaje ha aumentado recientemente, porque hoy las mujeres retrasan el casarse. Las cifras se refieren a los EE.UU.

para lograr compañía íntima o gratificación sexual! Dada la creciente aceptación del cohabitar entre solteros (y dado lo fácil del divorcio), ¿no parece efectivamente probable que el matrimonio tradicional llegue a ser todavía menos predominante?

Quizá la definitiva razón sociológica en pro del matrimonio sea la procreación y crianza de los hijos, pero el creciente desvío respecto a la paternidad debilita también un tanto este fundamento. De persistir las actuales tasas de primeros alumbramientos, cerca de un 30 por ciento de las mujeres estadounidenses que están ahora en edad apta para tener hijos no tendrían ninguno. La más alta proporción de las anteriormente registradas fue de un 22 por ciento de mujeres sin hijos entre las nacidas en los Estados Unidos en 1908. Parece probable que las jóvenes de hoy superarán aquel récord en unos cuantos puntos de porcentaje, aunque no lleguen al 30 por ciento que implican las actuales tasas. Si el gran aumento de nacidos fruto de parejas no casadas que se ha registrado en Suecia en los últimos años, a consecuencia del proporcional aumento de parejas que cohabitan sin matrimoniar, es un indicio de que se está debilitando el tabú que pesaba sobre la fertilidad extramarital, entonces hasta esta función del matrimonio puede que pierda importancia.

El declinar de la fertilidad cabe, pues, considerarlo como una causa y a la vez una consecuencia de cambios en el matrimonio y en la familia. El hecho de tener pocos hijos o ninguno se puede interpretar como algo que facilita la actividad económica de las mujeres e incrementa su igualdad económica con los hombres, dos desarrollos que hacen, a su vez, que el matrimonio y la crianza de hijos tengan menos de respuesta social automática. Un futuro más incompatible cada día con la concepción del duradero matrimonio tradicional será, casi con certeza, un futuro de bajo nivel de fertilidad en los países desarrollados.



TASA DE MATRIMONIOS: Ha ido generalmente descendiendo en Suecia *(trazo negro)* y Dinamarca *(trazo gris)* desde 1965, y en los Estados Unidos *(trazo de color)* desde 1973. Cada vez hay más parejas que hacen vida marital. En Suecia se ha registrado un gran aumento de hijos nacidos de parejas no casadas.

Hay una amplia gama de opiniones acerca de cuál sea la tasa óptima del crecimiento de población en el mundo desarrollado. Arguyen algunos coherentemente que el crecimiento ha excedido ya del nivel que imponen consideraciones tocantes al medio ambiente y a la calidad de la vida, y que la población adicional que está ya "a las puertas" agudizará el deterioro de la situación. Desde este punto de vista, es deseable una futura merma real del número de habitantes. Otros han llegado a la conclusión de que el mejor resultado sería el cero de aumento de población. El Comité para el estudio del Aumento de Población y del Futuro de América examinó las implicaciones económicas, ambientales y gubernamentales de varias tasas de crecimiento y concluyó, en 1972, que cualquier crecimiento que rebasara aquel al que los Estados Unidos estaban ya comprometidos por el legado de su crecimiento anterior no haría más que agravar los problemas de la nación y dificultar más su solución.

No hay, sin embargo, ninguna cualidad mágica en la tasa de la fertilidad 2,1 que mantiene precisamente el nivel de remplazo. Nadie ha inventado todavía un mecanismo homeostático que regule automáticamente la reproducción de una sociedad para mantenerla a un adecuado ritmo. Si la fertilidad sigue bajando, entonces la prognosis definitiva ha de ser: crecimiento de población negativo, con el consiguiente descenso del número de habitantes y envejecimiento de las poblaciones. Merece la pena insistir en que por los años 30 se habían aventurado algunas predicciones en este sentido, las cuales resultaron después ridículas a la luz de la explosión demográfica que las siguió, y esto podría suceder de nuevo. Lo que yo sostengo es que (si mi interpretación de los datos es, incluso sólo aproximadamente, correcta) un resurgir importante de la fertilidad parece improbable y que el mundo desarrollado ha reemprendido en cuanto a ésta el largo camino de la bajada que comenzó a seguir hace un siglo o más.

Lo que está causando preocupación, especialmente en Europa, donde el proceso entero del cambio está más avanzado que en los Estados Unidos, es que la fertilidad vaya a caer bastante por debajo del nivel de remplazamiento, que la proporción de viejos aumente por ello mucho, que la población no se mantenga, por consiguiente, estacionaria sino que empiece a disminuir... y que al aumentar en edad a esa población se le haga cada día más difícil retrasar el momento en que comience el declinar demográfico.

Varios países han dado ya señales de sentirse muy incómodos con su tendencia a disminuir en fertilidad. Un reciente análisis de las Naciones Unidas revelaba que 10 países europeos consideran muy baja su actual tasa de fertilidad. En Alemania Occidental (donde en 1976 hubo 733.000 óbitos y pocos más de 603.000 nacimien-

tos), la creciente preocupación se reflejó recientemente en un programa televisivo titulado "¿Se extingue el pueblo alemán?", pregunta a la que se contestaba en sentido afirmativo. Un demógrafo alemán ha dicho: "Por debajo de las apariencias somos un pueblo que se extingue".

En Francia, donde el aumento de población ha sido durante mucho tiempo objeto de interés tanto por cuestión de prestigio nacional como por razones específicas de seguridad de la nación, este asunto ha llegado a someterse a un debate nacional. En 1976, el Presidente Giscard d'Estaing advirtió que ningún país sólo medianamente poblado puede aspirar con realismo a ser grande. El que fue primer ministro, Michel Debré, extremista en este tema, considera un peligro para la nación el descenso de la tasa de nacimientos (que, siendo en 1976 del 13,6, aún estaba por encima de la de defunciones, que era del 10,5). Hasta llegó a proponer una reforma de la ley electoral por la que se otorgase a los padres un voto más por cada hijo que tuviesen de menor edad que la precisa para contar con el derecho al voto. Durante muchos años el gobierno francés ha venido ayudando a las familias numerosas con subsidios como concesión de créditos especiales y exoneración de impuestos. Desde 1976 garantiza también unos ingresos mínimos a las madres solteras, un permiso de dos años por maternidad a las empleadas en los servicios civiles y la exención del servicio militar a los padres jóvenes que se hallen en especiales circunstancias.

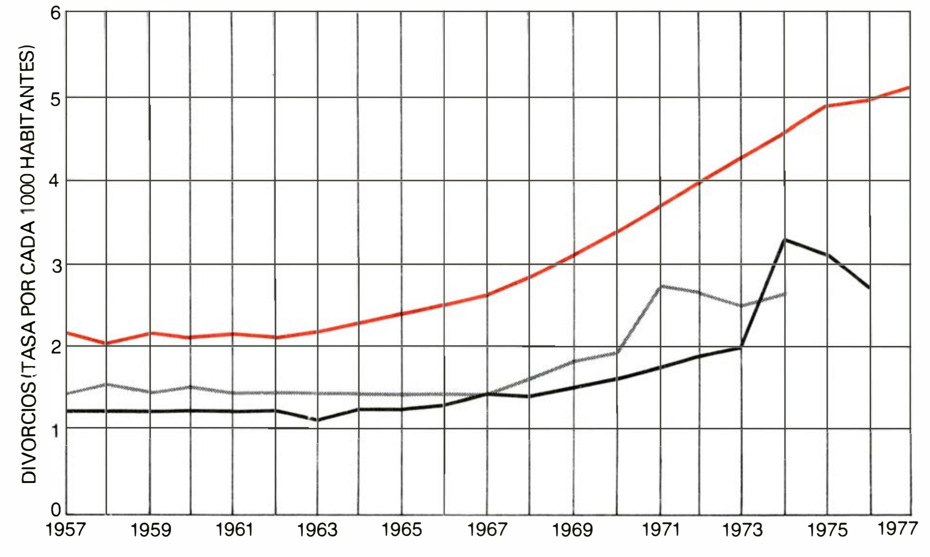
Muchas de las medidas de bienestar social tomadas en Suecia se han concentrado desde hace mucho en la protección de la infancia y la asistencia a los padres (estén o no casados). Un ejemplo reciente es el de la ley del "seguro de los padres", que aumenta de seis meses a nueve el período de vacaciones pagadas que le corresponde a la madre o al padre de un nuevo bebé; período de vacaciones que se puede tomar dentro de un plazo de ocho años.

Alemania Oriental, donde la tasa de nacimientos había bajado, en 1975, hasta el 10,6 (siendo la de muertes el 13,5), logró, según las noticias, elevar la tasa de nacimientos al 13,3, en 1977, gracias a la introducción de varias medidas nuevas a mediados de 1977. Incluíanse entre ellas la reducción de la semana laboral para las madres de dos o más hijos, la ampliación del permiso pagado por causa de embarazo de 18 a 26 semanas y varios incentivos financieros, el más notable de los cuales es un crédito sin interés de hasta 10.000 dólares para la adquisición del piso y del ajuar. La deuda se reduce en 1000 dólares si nace un hijo antes de los ocho años del casorio, y en otros 1500 dólares si nace un segundo hijo, siendo eliminada por completo al nacer un tercer hijo. El costo de estas ayudas es considerable: se habla de 37 millones de marcos en créditos, además de las horas de trabajo que esas mujeres no efectuarán. Según un informe, "90.000 mujeres –el 1 por ciento del total de la fuerza de trabajo– han aceptado la invitación del gobierno a que se dediquen a producir bebés".

En otros países de la Europa del Este se han desplegado activas campañas en pro de la natalidad. Rumanía logró cierta notoriedad entre los especialistas en política del poblamiento con su demostración del impacto demográfico que produce la prohibición legal del aborto: prohibido éste en 1966, la tasa de nacimientos subió a más del doble en 1967. (De entonces acá ha ido bajando poco a poco, a medida que el aborto ilegal y los métodos anticonceptivos han sido sustituidos por el aborto legal.) Las leyes del aborto se han endurecido también en estos últimos años en Bulgaria, Checoslovaquia y Hungría, con lo que han aumentado las tasas de nacimientos. Hasta qué punto sea atribuible tal aumento a las medidas económicas pronatalistas que han sido introducidas a la vez (como permisos por maternidad, créditos familiares, servicios subsidiarios por nacimiento del hijo, alojamientos preferentes y otros estímulos), más que el cambio de las leyes del aborto, es cuestión por estudiar.

Tales respuestas al declinar de la fertilidad sugieren que irá en aumento el número de gobiernos que consideren la baja fertilidad y el crecimiento negativo de la población como una amenaza y que desarrollen programas para inducir a las mujeres a casarse y a tener más hijos que los que hoy parecen motivadas a tener. ¿Qué tipo de medidas será el adecuado para lograr este objetivo? La misma pregunta se estaba formulando hace 40 años cuando varios países europeos se veían ante parecidos futuros demográficos (perspectiva que fue radicalmente alterada por la Segunda Guerra Mundial y sus secuelas). Por aquel entonces se especuló que la sociedad podría tener que desarrollar una clase de especialistas en reproducción, a los que se compensaría por sus servicios y cuya progenie sería criada en centros especializados, sin las limitaciones que comporta la estructura familiar convencional. Semejantes planes fueron desbaratados por la avalancha hacia el matrimonio, los suburbios y la cría de hijos que siguió a la Guerra Mundial, y todavía pertenecen a la esfera de la ciencia-ficción.

Dadas las actuales tendencias no es difícil imaginarse una sociedad en la que un tercio de las mujeres no tengan hijos, lo que significaría que los dos tercios restantes habrían de tener un promedio ligeramente superior al de tres hijos cada una para remplazar a su generación. No parece probable que los flojos y a menudo triviales incentivos puestos hasta ahora en juego por los gobiernos logren motivar tanto las funciones genésicas que se alcance el nivel de fertilidad requerido. Cuesta imaginarse que a unas mujeres bien educadas y bien pagadas, que han crecido en una sociedad que recompensa las dedicaciones no parentales, les atraigan unos pocos miles de dólares de beneficio lo bastante como para renunciar a sus ingresos ordinarios, interrumpir una carrera y dedicar el tiempo y la energía necesarios para criar a varios hijos. La nación industrializada que decida no tolerar una tasa de crecimiento negativa, quizá tenga que subvencionar a gran escala la reproducción, con una inversión muy importante de sus fondos públicos. Casi con



TASA DE DIVORCIOS: En los Estados Unidos *(trazo de color)* es elevada y ha venido aumentando; en Suecia *(trazo negro)* y en Dinamarca *(trazo gris)* es más baja que en Norteamérica, pero se ha ido elevando.

toda certeza tendría que garantizarles a las mujeres que trabajan el disfrute de amplios permisos y vacaciones pagadas para atender a la crianza de sus hijos, así como que no sufrirían por ello pérdidas de ingreso o de posición, y que podrían contar con la ayuda de un amplio y bien dotado sistema de guarderías y demás instituciones para el cuidado de los niños.

Conforme se acerca más en su bajada la tasa de nacimientos de un país a la tasa de defunciones, la inmigración desempeña un papel creciente como determinante del aumento de población. En teoría, el gobierno que tratara de mantener fija una tasa de aumento –por ejemplo el cero de crecimiento– podría manejar sus cuotas de inmigración para hacer que ésta se igualase poco más o menos al déficit del número de nacidos en el año procedente. Las restricciones a la inmigración podrían reglamentarse de modo que no se impidiera entrar a sujetos bien cualificados con saberes y técnicas de los que hubiese necesidad en el país, siempre y cuando la economía pudiera mantenerse lo bastante fuerte para atraer a tales individuos. Claro que conjugar con acierto todos estos factores suele ser difícil. Los inmigrantes tienen edades distintas y, lo que puede llegar a ser más grave, difieren también por sus colores y nacionalidades y hablan diversas lenguas. Todo país con grandes contingentes de inmigración se ha de enfrentar más tarde o más temprano con problemas que se derivan de las diferencias en las lenguas, las costumbres, la religión o la raza y que tienden a contrarrestar muchas de las ventajas económicas y culturales a largo plazo.

En las últimas dos décadas, poco más o menos, la mayoría de los países de la Europa occidental y nórdica han admitido inmigrantes (o en muchos casos se han limitado a admitir trabajadores temporeros) procedentes de países menos ricos, con su población en aumento y con alto nivel de desempleo. La afluencia de inmigrantes ha proporcionado una fuerza de trabajo adicional, sobre todo para quehaceres relativamente bajos, que ha contribuido a mantener el rápido desarrollo económico, pero también ha creado problemas sociales. Los países en que escasea la mano de obra seguirán recurriendo, sin duda alguna, a la importación de trabajadores extranjeros mientras haya un excedente de mano de obra aprovechable en los países vecinos en los que sea alto el nivel de fertilidad. En cambio, ningún país ha optado aún por la inmigración por razones puramente demográficas, y la experiencia adquirida con los trabajadores extranjeros es tal que no es probable que los países receptores la olviden al enfrentarse con el más básico problema del complementar el aumento de población.

En los Estados Unidos preocupa cada vez más la inmigración ilegal, aunque no se tiene una idea clara de su magnitud ni de su duración. Sus efectos son locales, pero indudablemente ejerce un importante influjo económico en diversas partes del país. De hecho, lo que hace que la inmigración ilegal persista es su valor económico; si no fuese tan valioso para las localidades y para los empresarios particulares, el aflujo ilegal resultaría presumiblemente controlado con mayor eficacia por las autoridades estadounidenses.

Sean cuales fueren las ventajas económicas que a corto plazo reporta la migración internacional, legal o ilegal, lo que está claro es que no ofrece ninguna solución a largo plazo para un país cuya fertilidad en nacimientos esté bastante por debajo del nivel requerido para el adecuado remplazo de la población.

En los Estados Unidos, por ejemplo, el volumen de inmigración tendría que ser aproximadamente el doble del que totaliza la actual inmigración legal para evitar la eventualidad del así llamado "crecimiento negativo".

Los países desarrollados están empezando a ajustarse ahora a la previsión de una población estacionaria después de varios siglos de aumento. Si el cero de aumento de población acaba por quedar atrás y el "crecimiento negativo" parece ser el curso del futuro, pueden esperarse revolucionarios cambios sociales en el siglo XXI. Sabemos muy poco acerca de las consecuencias que a corto plazo pueda tener el "aumento negativo" de población, pero la consecuencia a largo plazo es bastante obvia.

Debo aclarar que aquí no estoy tratando de urgir como deseable el mantenimiento de una determinada tasa de crecimiento. El Comité para el estudio del Aumento de Población y del Futuro de América concluía, tras una exhaustiva evaluación de los pros y los contras del aumento poblacional, que el lento y eventual cese de aumento parece ser lo que más interesa a los Estados Unidos y que incluso breves períodos de "crecimiento negativo" pueden serles beneficiosos. Lo cierto es que el cero de crecimiento de población está ahora a la vista en muchos países desarrollados. Si la población disminuye durante algún intervalo de tiempo considerable, esta merma será vista –se acierte o no en ello– con alarma. Antes de lanzarse a desarrollar de un modo acrítico políticas de estímulo a la fertilidad, sería conveniente y deseable que se hicieran serios esfuerzos por evaluar los costes y beneficios de las diversas alternativas posibles en cuanto a las tasas del aumento de población.

Moléculas nucleares

A ciertas energías, los protones y neutrones de los núcleos forman racimos, que se mueven los unos con respecto a los otros en órbitas sencillas. Estas moléculas nucleares constituyen una nueva clase de dinámica y estructura del núcleo

D. Allan Bromley

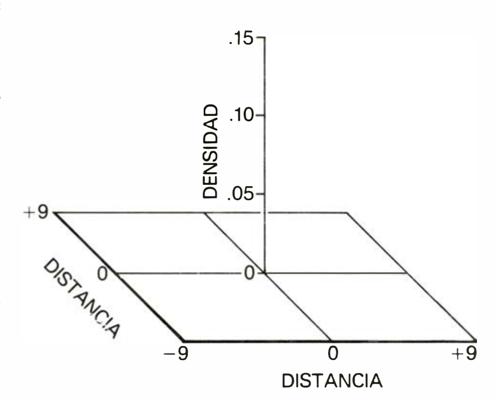
La tendencia de los átomos a unirse en moléculas por el hecho de intercambiar o compartir los electrones más externos, o de valencia, es un fenómeno bien conocido. En el centro de los átomos están los núcleos atómicos, que son aproximadamente 100.000 veces más pequeños. Se ha demostrado que también los núcleos pueden formar moléculas: bajo condiciones especiales en interacciones de alta energía, quedan momentáneamente ligados por enlaces efectivos. Si estos enlaces son debidos o no a que los núcleos intercambian o comparten los protones o neutrones de valencia aún no está claro. Estas moléculas nucleares de vida media muy corta constituyen un modo enteramente nuevo de estructura y dinámica nuclear, que sólo ahora los físicos empiezan a entender.

La estructura y estabilidad de las moléculas atómicas y las moléculas nucleares dependen de las características de las fuerzas fundamentales de la naturaleza que rigen las interacciones entre electrones, protones y neutrones. Las moléculas atómicas son el resultado del balance de dos aspectos de largo alcance de las fuerzas electromagnéticas: la atracción covalente de las nubes electrónicas de los átomos y la repulsión electrostática de los núcleos de los átomos, que están cargados positivamente. La distancia entre los átomos en una molécula atómica es tal que las fuerzas de atracción y repulsión se cancelan las unas a las otras. En general, las moléculas atómicas son relativamente estables y se mantienen unidas a no ser que sean rotas por alguna interacción energética. En las moléculas nucleares se alcanza un balance menos estable entre la misma repulsión elestrostática de largo alcance y la mucho más intensa interacción nuclear de corto alcance que determina los movimientos de los protones y neutrones. Las fuerzas nucleares no actúan sobre los electrones y, por tanto, no juegan ningún papel directo en la formación de las moléculas atómicas. En las moléculas nucleares, las fuerzas nucleares y electrostáticas son tan fuertes, comparadas con las fuerzas de las moléculas atómicas, que, en las moléculas nucleares, el balance de fuerzas es muy precario tendiendo a romperse bajo la repulsión electrostática, o a fundirse formando un núcleo bajo la atracción nuclear.

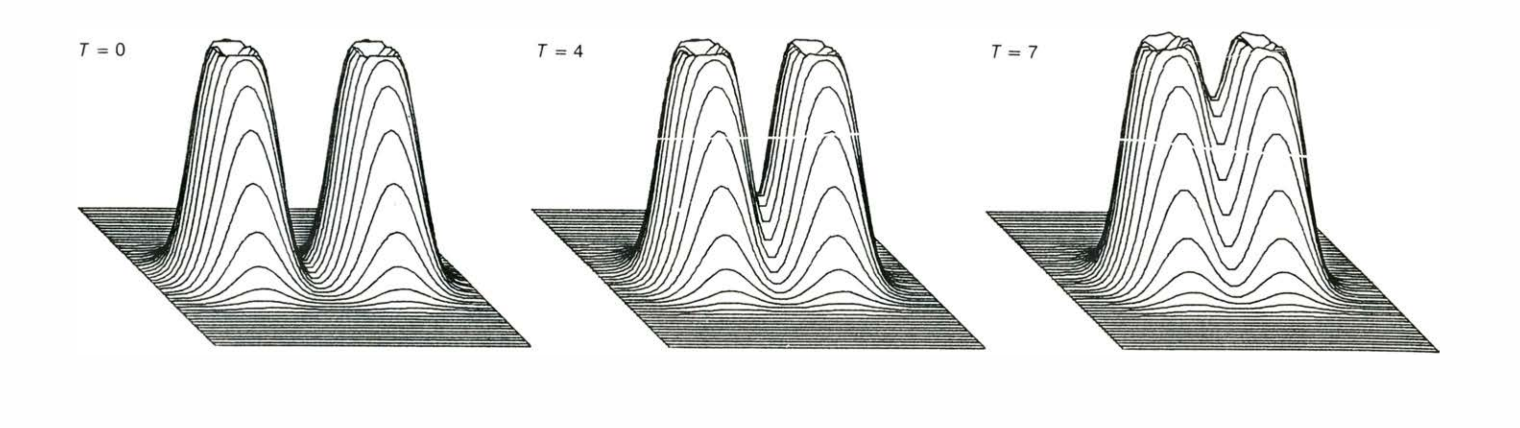
Si se pretende formar una molécula nuclear, hay que acercar dos núcleos lo suficiente para que las fuerzas nucleares de corto alcance puedan compensar la fuerza de repulsión electrostática de largo alcance. En la naturaleza, en general, los núcleos no están lo suficientemente cerca unos de otros para que las fuerzas nucleares atractivas de corto alcance puedan formar moléculas nucleares. Quizá puedan detectarse moléculas nucleares en los interiores de estrellas gigantes, donde una tercera fuerza fundamental, la gravedad, puede adquirir intensidad suficiente para contrarrestar la repulsión electrostática y empujar unos núcleos contra otros.

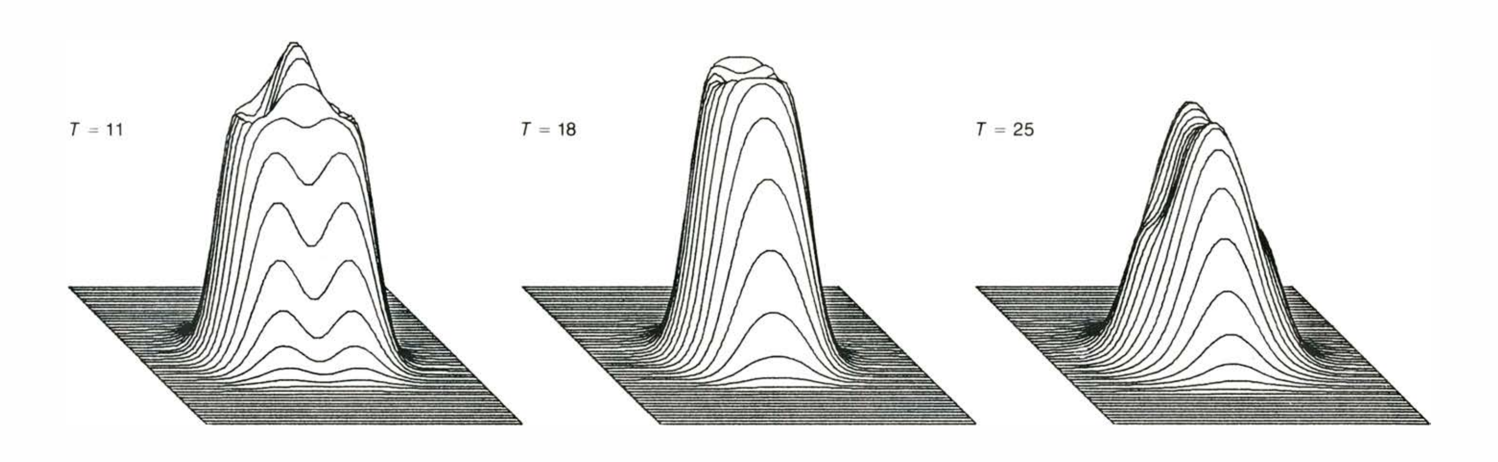
Si bien no se han encontrado moléculas nucleares en la naturaleza, ya han sido producidas en los laboratorios al hacer chocar núcleos muy acelerados con núcleos en reposo. Los aceleradores de partículas pueden elevar la energía cinética de un núcleo proyectil hasta el punto que puede colocarse muy cerca de un núcleo que actúa como blanco, venciendo la fuerza de repulsión electrostática. Entonces, bajo determinadas condiciones, parecen formarse moléculas nucleares; los dos núcleos vibran y giran rápidamente el uno con relación al otro durante un tiempo muy breve antes de dispersarse o de fundirse en uno solo.

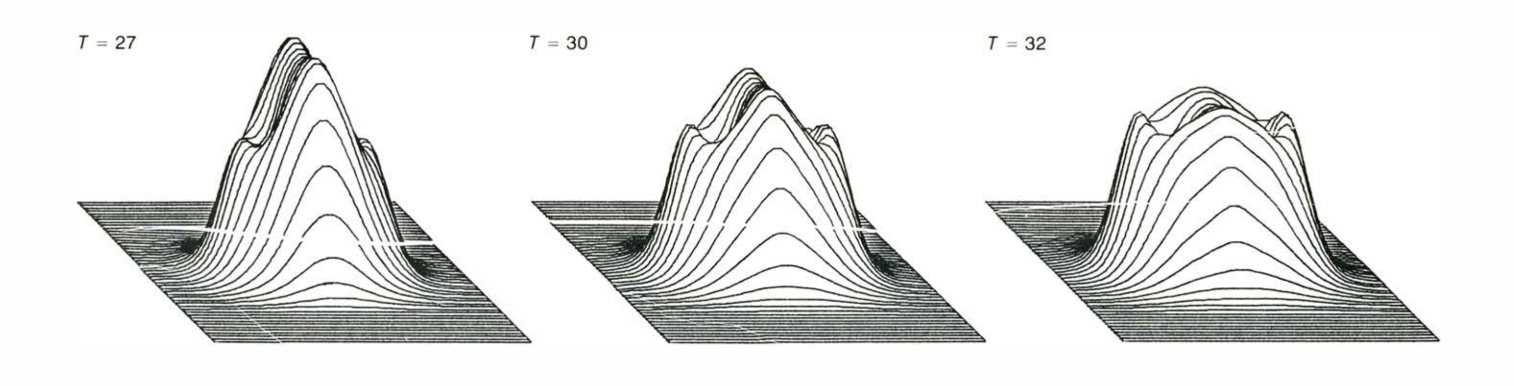
Las moléculas nucleares son entidades bien diferenciadas porque tienen una vida más larga (alrededor de 10^{-21} segundos) que el tiempo que tardan dos núcleos en chocar (unos 10^{-23} segundos), aunque no duran un tiempo suficientemente largo para poder someterlas a los mismos tipos de medición que pueden hacerse con moléculas atómicas. Pero podemos aprender mucho sobre ellas de forma indirecta, estudiando cómo su formación y desintegración afecta a la "sección eficaz" o probabilidad de que sucedan ciertas cosas cuando dos núcleos de energías específicas chocan y forman unos determinados productos nucleares finales. Allí donde la sección eficaz varía suavemente con la energía del núcleo proyectil, no tienen

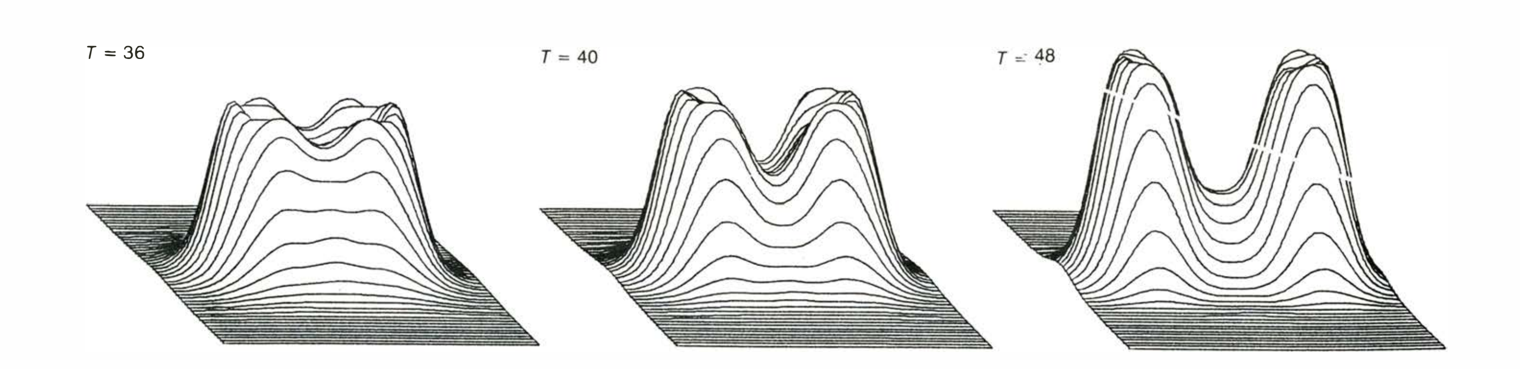


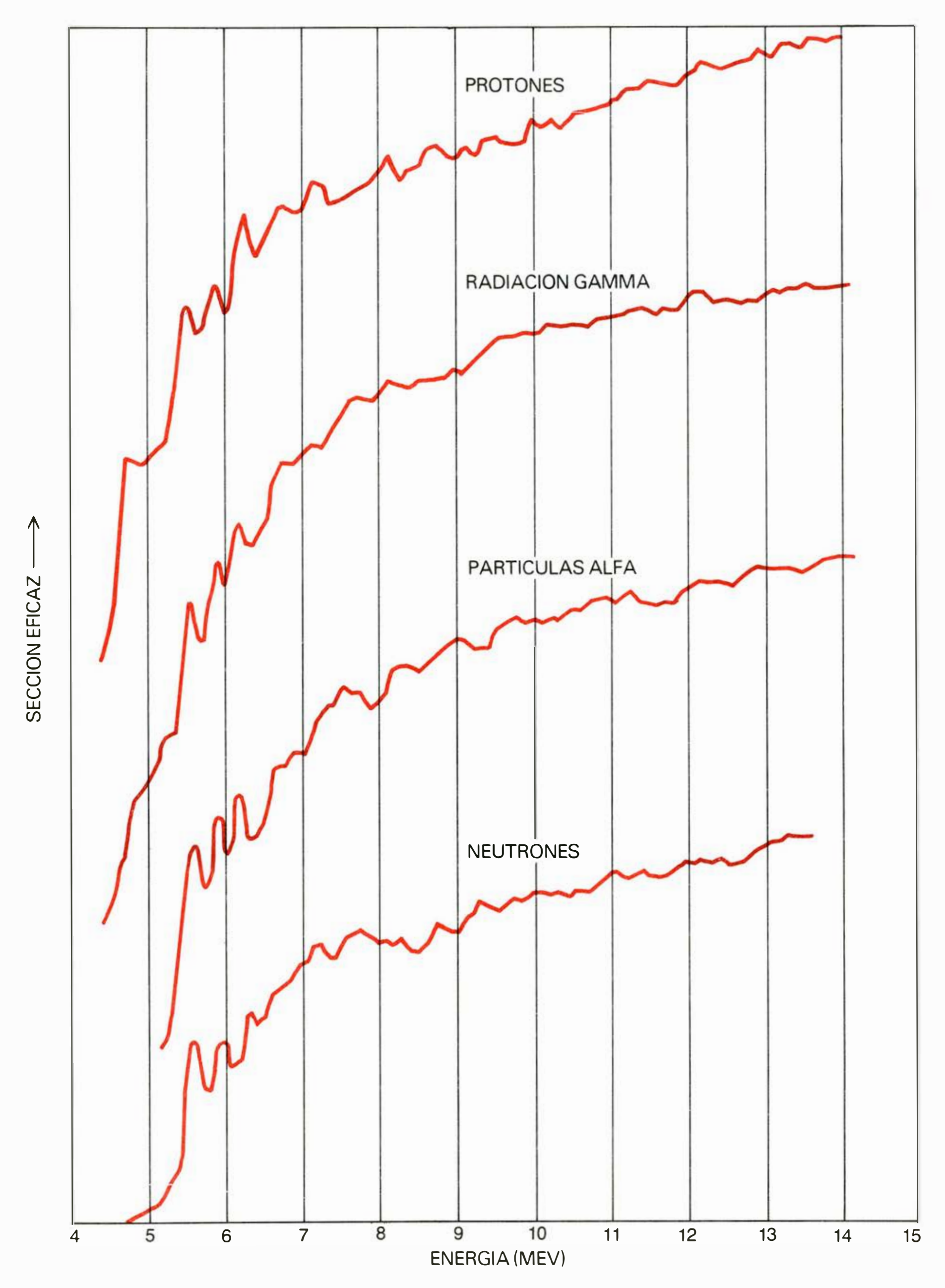
FORMACION de una molécula nuclear, representada en estos dibujos hechos con ordenador mediante la densidad calculada de protones y neutrones en dos núcleos de carbono 12 en colisión. La densidad viene dada en términos del número de protones y neutrones por 10^{-39} centímetros cúbicos. Cada protón y neutrón del proyectil de carbono que incide sobre un blanco de carbono tiene una energía media de 64 Mev (millón de electronvolt). La interacción entre los dos núcleos dura 5 x 10^{-22} segundos. La unidad de tiempo T es 3,3 x 10^{-24} segundos y la unidad de distancia es 10^{-13} centímetros. En $T = 0$, los núcleos de carbono entran en contacto y a $T = 11$ la densidad alcanza un máximo en la punta central. Se desarrolla luego un cuello de larga vida que une los dos fragmentos, cada uno de los cuales tiene aproximadamente la densidad de uno de los núcleos del carbono original. Por último, la molécula nuclear empieza a romperse, pero la capacidad limitada del ordenador hizo imposible llevar a cabo los cálculos hasta ver cómo los dos núcleos de carbono quedaban finalmente separados. El hecho de que los perfiles de densidad de cada núcleo sean menos suaves a $T = 48$ que a $T = 0$ significa que la energía había sido transferida del movimiento relativo hasta excitaciones internas. Los cálculos fueron realizados por Ronald Y. Cusson y Joachim Maruhn, del Oak Ridge Laboratory.


T = 0
T = 4
T = 7

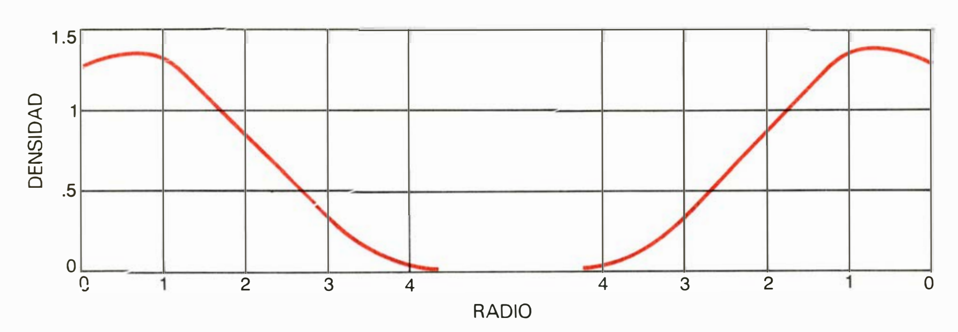

T = 11
T = 18
T = 25


T = 27
T = 30
T = 32


T = 36
T = 40
T = 48



SECCION EFICAZ, O PROBABILIDAD de que los distintos productos nucleares se han emitido de la colisión entre núcleos de carbono-12, representada en función de la energía cinética disponible. A una energía de 6,6 MeV, la barrera de Coulomb. los núcleos que chocan pueden sólo tocarse. Resultan de particular interés las tres resonancias o picos que aparecen exactamente por debajo de la barrera de Coulomb. Las tres resonancias, que aparecen en las cuatro secciones eficaces y para todos los ángulos de observación, constituyeron la primera prueba de que existían fenómenos de moléculas nucleares. Los datos fueron obtenidos en 1960 por el autor, Einar Almqvist y John A. Kuehner, de los Chalk River.



PERFILES DE DENSIDAD para dos núcleos de carbono-12 en colisión con una energía de 5 MeV, que se observan cuando los centros de los núcleos se colocan a una distancia mínima de 10.4 x 10^{-13} centímetros. Tal separación duplica, y más, no sólo el radio medio del núcleo de carbono, sino también el radio mayor efectivo. En esta ilustración, el radio viene medido en unidades de 10^{-13} centímetros.

lugar fenómenos de vida larga; los núcleos se encuentran en contacto un tiempo muy breve, como en una colisión elástica donde los dos núcleos se dispersan a imagen de lo que ocurre en el choque de dos bolas de billar.

Si, por el contrario, la sección eficaz resuena, o forma un pico, a ciertos valores de la energía, ello sugiere que los núcleos pueden permanecer próximos uno a otro durante un intervalo de tiempo relativamente largo. Y si otras condiciones, que discutiré más adelante, se satisfacen, estas resonancias se pueden identificar como signos de la formación de moléculas nucleares. Usando el principio de indeterminación de Werner Heisenberg, la vida media de tal molécula puede calcularse directamente a partir de la anchura de la correspondiente resonancia. La anchura representa la indeterminación en la energía de la molécula nuclear y el principio de indeterminación asocia una vida media larga con una resonancia estrecha (una pequeña indeterminación) y una vida media corta con una resonancia ancha (una gran indeterminación)

En 1960, Einar Almqvist, John A. Kuehner y yo en Chalk River Nuclear Laboratories, Canadá, descubrimos resonancias en colisiones entre núcleos de carbono-12 (^{12}C), que constituyeron la primera prueba de la existencia de moléculas nucleares. Usamos el primer acelerador electrostático Van de Graaff de tipo tándem para juntar los núcleos de carbono. A ciertas energías, observamos resonancias en la producción de ciertos subproductos del choque: protones, radiación gamma, partículas alfa y neutrones. Las resonancias tenían anchuras características de unos 120 KeV (mil electronvolt), que de acuerdo con el principio de indeterminación correspondía a una vida media de 5×10^{-21} segundos, un valor más de 10 veces mayor que la duración típica de un simple choque. Sin embargo, no se podía concluir sólo a base de este tiempo de colisión, comparativamente largo, que las resonancias habían sido producidas por la formación de moléculas nucleares y no por algún otro fenómeno.

Pero encontramos una corroboración adicional de que habíamos detectado moléculas nucleares. En las moléculas atómicas cada átomo retiene la mayor parte de sus propiedades internas, a pesar del hecho de que todos los átomos estén ligados entre sí. De la misma forma, en las moléculas nucleares debíamos esperar que cada núcleo constituyente retuviera la mayoría de sus propiedades internas. En efecto, esto fue lo

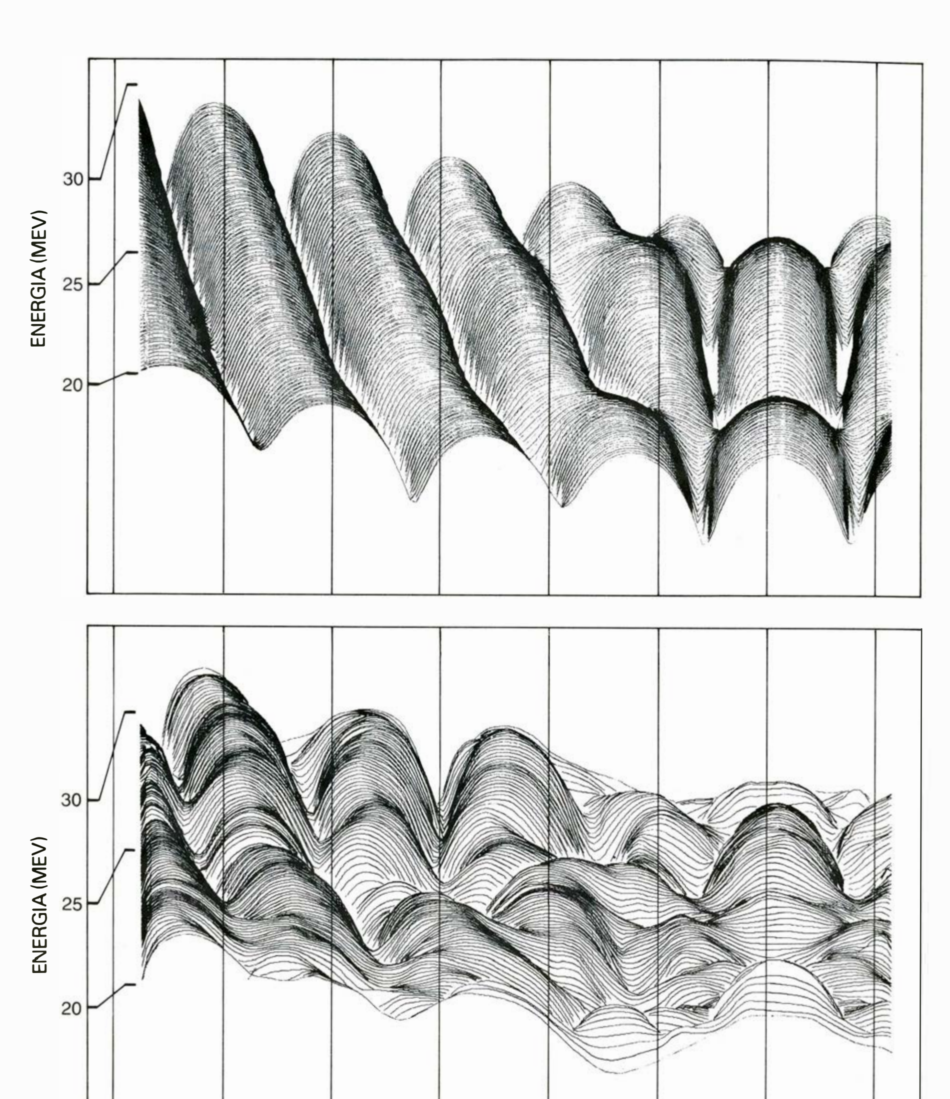
que sucedió en el caso de nuestro experimento en Chalk River.

Hicimos mediciones detalladas que demostraban que, a las energías de resonancia, la probabilidad del choque que diera como resultado dos núcleos de carbono-12 era 25 veces mayor que la que se habría obtenido si los dos núcleos se hubieran fundido momentáneamente para formar un núcleo de magnesio-24 normal (^{24}Mg), que tiene el mismo número de protones y de neutrones que dos núcleos de carbono-12. Esto aportaba gran información sobre la dinámica de las colisiones del carbono-12 y la estructura del magnesio-24. Nuestras mediciones de secciones eficaces indicaban que, a las energías de resonancia, 24 protones y neutrones preferían formar dos núcleos de carbono-12 ligeramente ligados en lugar de un núcleo normal de magnesio-24. Por esta razón, sugerimos que, si a un núcleo de magnesio-24 se le suministraba por cualquier medio una cantidad de energía correspondiente a las resonancias observadas, habría una gran probabilidad de que sus 12 protones y sus 12 neutrones se reordenaran en dos núcleos de ^{12}C, formado cada núcleo por 6 protones y 6 neutrones.

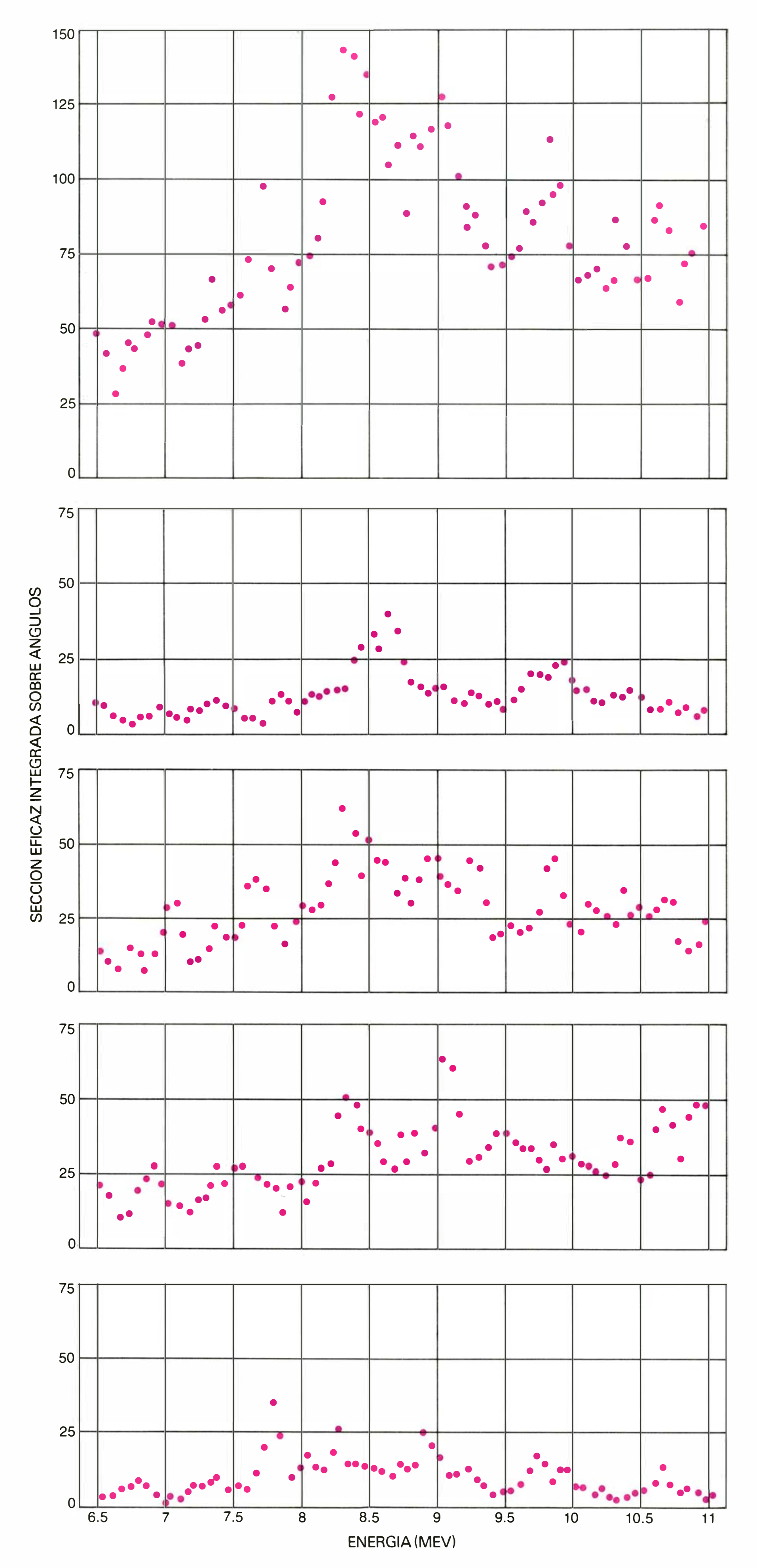
Otros físicos nucleares intentaron comprobar esta sugerencia. En la primera prueba, llevada a cabo por N. O. Lassen y colaboradores, de la Universidad de Copenhague, se creó magnesio-24 cuando se unían un núcleo de neon-20 (10 protones y 10 neutrones) y un núcleo de helio-4 (dos protones y dos neutrones). Al ajustar la energía total de manera que fuera la misma que la de las resonancias de Chalk River, aparecía una gran abundancia de fragmentos de carbono-12. En otro experimento realizado por Norman K. Sherman, de la Universidad de la Reina, de Canadá, se hacía que un núcleo de magnesio-24 absorbiera fotones de alta energía de suerte que su energía se elevaba hasta los valores de resonancia, en cuyo instante el núcleo de magnesio-24 se fisionaba con un rendimiento aumentado de carbono-12, lo que constituye la señal de una molécula nuclear.

Dejando de lado el papel de las moléculas nucleares como nuevos modos no esperados de estructura y dinámica nucleares, aquéllas han interesado a los físicos por el papel que podrían jugar en ciertos fenómenos astronómicos, la evolución estelar por ejemplo. La combustión de carbono nuclear se ha considerado, desde hace tiempo, el proceso crucial en el período final de la existencia de una estrella. Obviamente, la presencia de resonancias moleculares no esperadas en la interacción nuclear ^{12}C-^{12}C afectaría el ritmo a que se quema el carbono. A principios de la década de los años setenta, investigadores del Instituto de Tecnología de California y de la Universidad de Pennsylvania extendieron nuestras mediciones de Chalk River a energías menores, más próximas, presumiblemente, a aquellas que caracterizan el quemado de carbono en las estrellas. Encontraron nuevas resonancias a energías del carbono disponibles entre 3 y 4 MeV (millón de electronvolt); extrapolaron la sección eficaz hasta energías de unos 100 KeV, que corresponden a las energías que los núcleos alcanzarían en el interior de estrellas con temperaturas de unos 1200 millones de grados Kelvin, donde se supone que tiene el lugar el quemado del carbono. La extrapolación a energías astrofísicas condujo a una hipótesis intrigante: el quemado de carbono procede a un ritmo que es más de 10 veces más alto que el permitido por cualquier modelo que se haya propuesto para explicar la evolución estelar.

Hubo otras sorpresas en las primeras mediciones. En nuestro trabajo original encontramos que no aparecían resonancias análogas en las colisiones de núcleos de oxígeno-16 (^{16}O), pero sí aparecían en colisiones entre núcleos de oxígeno-16 y carbono-12. Otros investigadores observaron pronto el mismo hecho. ¿Por qué aparecían claramente resonancias moleculares en las interacciones ^{12}C-^{12}C y ^{12}C-^{16}O (y recientemente en colisiones entre boro-11 y carbono-12 y entre carbono-12 y carbono-13), y no surgían en absoluto en la interacción ^{16}O-^{16}O? La pregunta quedó sin contestar a lo largo de la década de los sesenta y primeros años de la década de los setenta. Nos intrigaba saber por qué entre todas las clases de átomos, los de car-



MODELO MATEMATICO de tipo óptico. Incorpora fuerzas que actúan sobre los dos núcleos de carbono-12 cuando éstos se juntan y luego se dispersan elásticamente. Se usó para generar la superficie de la sección eficaz *(arriba)*. Los parámetros del modelo fueron ajustados hasta que la superficie calculada para la sección eficaz reprodujera la superficie de la sección eficaz medida *(abajo)*.



bono fueran los que se hallaran en número mayor en las moléculas atómicas. Pero no acertábamos a descubrir ninguna conexión entre el fenómeno atómico y el nuclear.

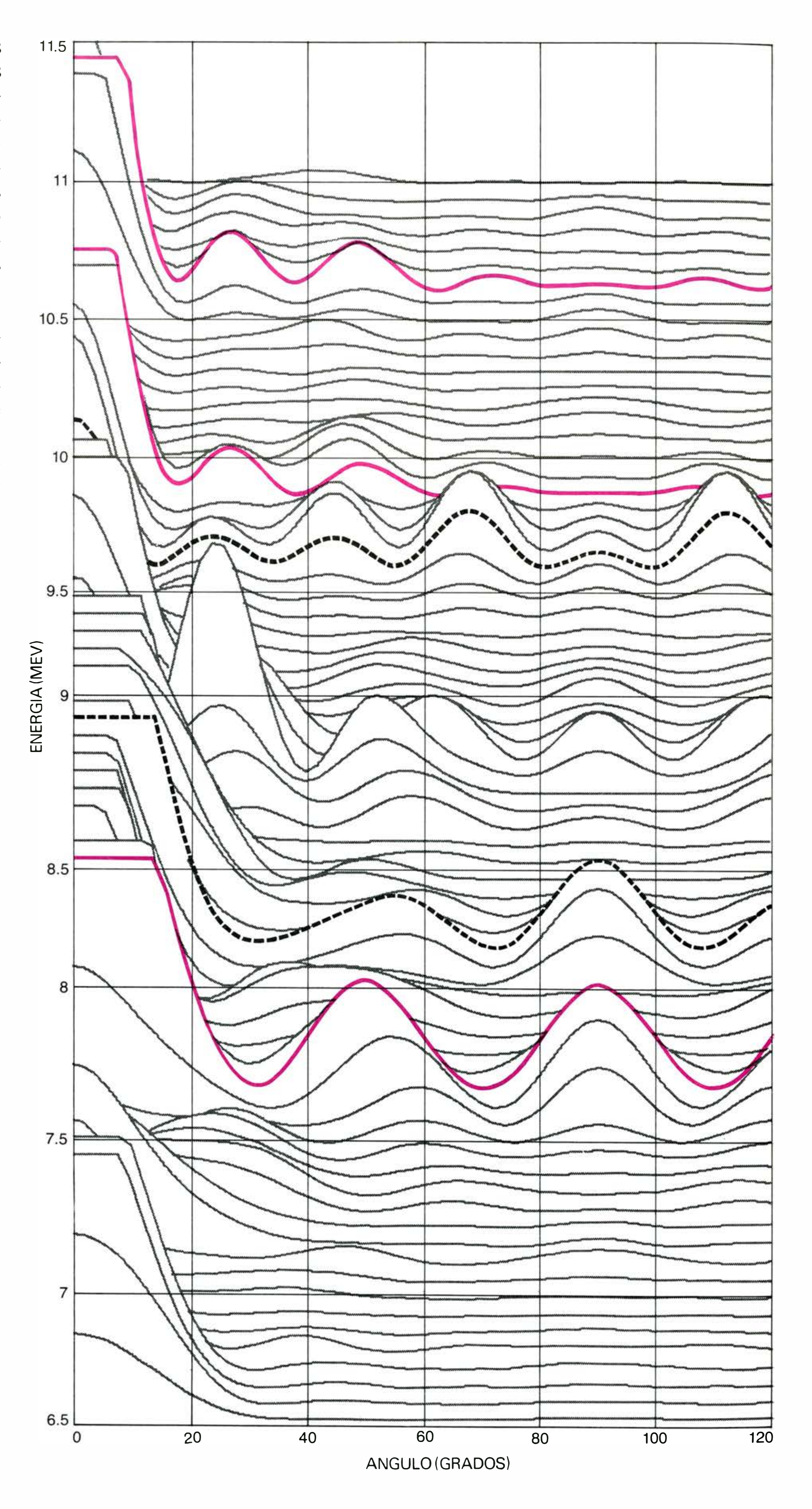
Los resultados originales de Chalk River estimularon un amplio interés experimental y teórico por las moléculas nucleares y otras interacciones de iones pesados (átomos que han sido ionizados de forma que sus núcleos pueden ser electrostáticamente acelerados). Aún así, el progreso fue muy limitado, por el lado experimental en razón de las energías, comparativamente bajas, de los haces de iones pesados de que se disponía por aquel entonces y, por el lado teórico, debido a la falta de explicación del singular papel que desempeñaba el carbono en las interacciones nucleares. Por lo que respecta al carbono, el modelo teórico de mayor éxito, desarrollado por Bunryu Imanishi, de la Universidad de Tokyo, podía reproducir los datos de Chalk River únicamente si se hacía lo que en aquellos tiempos se consideraba la extravagante hipótesis de que las superficies nucleares eran altamente transparentes: es decir, que las regiones de la superficie nuclear podían atravesarse unas a otras y salir sin ninguna perturbación reseñable. Más aún, el modelo de Imanishi pareció que quedaba desacreditado cuando falló a la hora de explicar las resonancias de energía más baja que habían sido descubiertas por los investigadores del proceso de combustión del carbono estelar. Todo lo cual trajo por consecuencia que los físicos empezaran a poner en duda la existencia de moléculas nucleares. Así, el trabajo en este campo disminuyó hasta el final de la década de los años 60, cuando nuevos y más potentes aceleradores de iones pesados entraron en funcionamiento.

Contemporáneo con la nueva tecnología, se registró un progreso en la comprensión de la posible estructura de las moléculas nucleares y de las condiciones en las que podrían aparecer. Aunque podrían esperarse que todas las moléculas nucleares surgieran, en condiciones apropiadas, como resonancias en las secciones eficaces de interacción, no todos los picos de las secciones eficaces

NEON 20 Y HELIO 4 (una partícula alfa) pueden formarse cuando colisionan dos núcleos de carbono-12 (^{12}C). Aquí se representa la sección eficaz del neon-20, integrada sobre todos los ángulos de dispersión, en función de la energía de los carbonos que colisionan. Los 4 dibujos inferiores corresponden a la formación de los 4 estados cuánticos de menor energía del neon-20. Algunas resonancias del dibujo superior, obtenido sumando las otras representaciones, corresponden a la formación de moléculas ^{12}C-^{12}C.

pueden identificarse como verdaderas resonancias y no todas las verdaderas resonancias corresponden a la formación de moléculas nucleares. Para identificar correctamente las moléculas, debemos ser capaces de reconocerlas en presencia de picos que caracterizaron otros fenómenos. El problema fue abordado por mi grupo de investigación que trabajaba con el primero de los grandes aceleradores Van de Graaff de tipo tándem, modelo "Emperor", instalado en el A.W. Wright Nuclear Structure Laboratory, de la Universidad de Yale. Comenzamos por desarrollar un modelo matemático de los rasgos más importantes de la colisión ^{12}C - ^{12}C; suponíamos en él que los núcleos se dispersaban por un potencial efectivo. Se trataba de una representación matemática de cómo la energía cinética de los núcleos que se aproximaban decrecía al disminuir la distancia que los separaba. El potencial incorporaba todas las fuerzas que actuaban sobre los núcleos cuando se aproximaban mucho: la repulsión electrostática, la atracción de las fuerzas nucleares y la fuerza centrífuga generada por la rotación de los núcleos alrededor de su centro de masas. Tal modelo se llama un método óptico porque usa el formalismo matemático que se desarrolló inicialmente para describir cómo la luz se dispersa por una bola de cristal obscuro. No obstante el destacado papel que los modelos ópticos desempeñan en física nuclear, atómica y de partículas, suelen usarse mal con bastante frecuencia. Se pretende que un tal modelo represente sólo las características medias de una interacción y, por tanto, no debe esperarse que reproduzca con precisión ningún dato específico, del siguiente tenor, por ejemplo: cómo varía el ritmo de aparición de un determinado producto a un cierto ángulo con la energía, o la forma en que el ritmo a una cierta energía varía con el ángulo. Con nuestro modelo generábamos superficies de secciones eficaces que dependían tanto de la energía como del ángulo [*véase la ilustración de la página 17*]. Las superficies muestran oscilaciones características en el ritmo de producción predicho para los núcleos dispersados cuando la energía, el ángulo o ambos variaban. Fuimos ajustando los parámetros de nuestro modelo hasta que pudiera reproducir las características más destacadas de nuestros datos experimentales con razonable precisión.

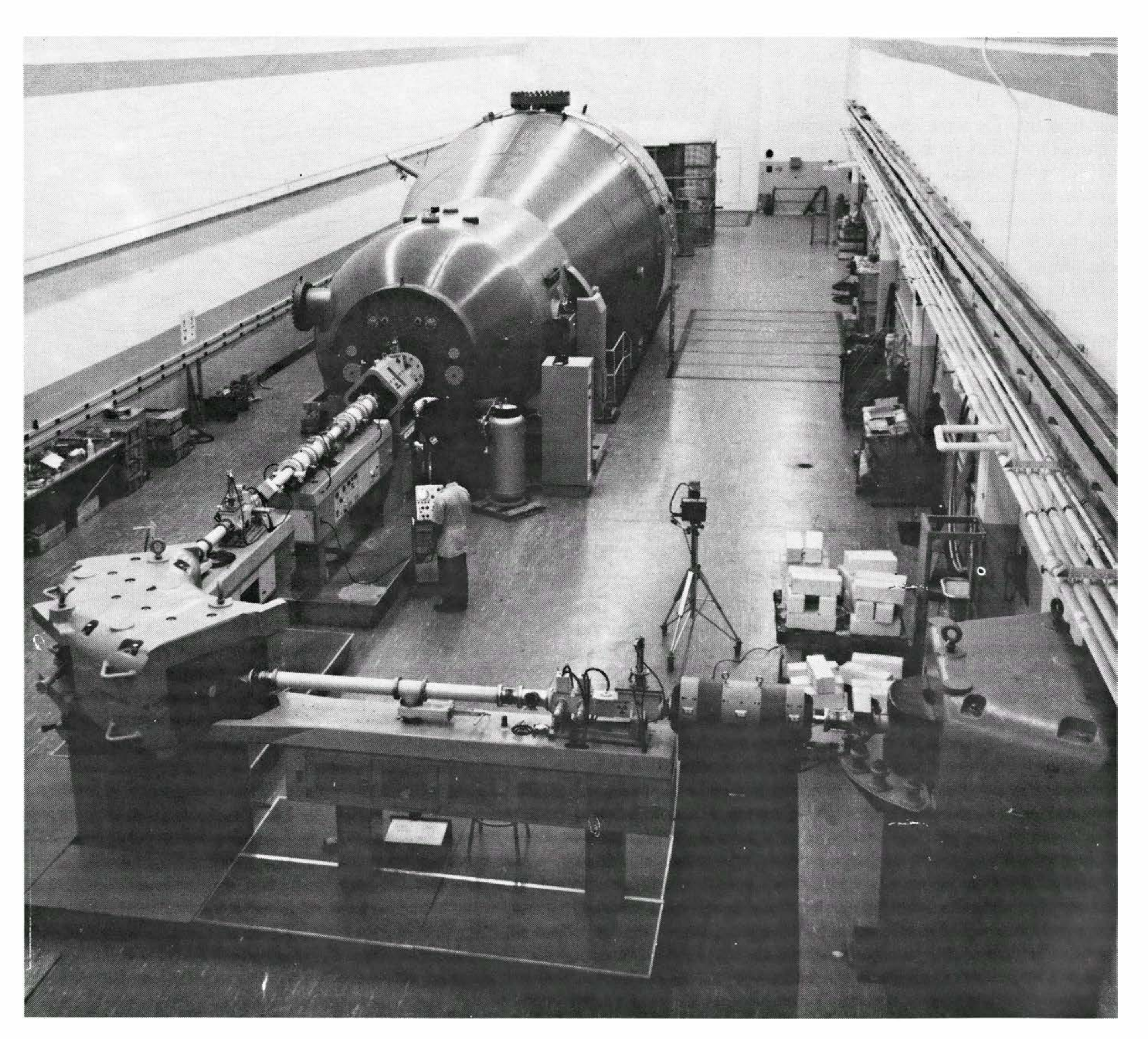
Sorprende, de un modo particular, en los nuevos datos de colisiones ^{12}C-^{12}C obtenidos por nosotros y por otros investigadores que aparezcan picos



PROBLEMA DE LAS RESONANCIAS ESPUREAS, ilustrado mediante el comportamiento de la superficie de la sección eficaz, medido para la creación de neon-20 en su estado fundamental, o estado de menor energía, en colisiones entre núcleos de carbono-12. Las curvas coloreadas fueron identificadas como resonancias que correspondían a verdaderas moléculas nucleares. Las curvas punteadas, aunque aparentemente equivalentes, se probó que estaban producidas por fluctuaciones estadísticas, y que se trataban de señales que no indicaban la formación de moléculas nucleares. (Dan Todd.)

muy estrechos en las secciones eficaces a lo largo de todo el intervalo de energías que han cubierto hasta ahora los experimentos. A las bajas energías estudiadas en Chalk River, Instituto de Tecnología de California, y Pennsylvania cuesta detectar las resonancias moleculares porque la repulsión electrostática entre los dos núcleos es tan fuerte que enmascara los efectos nucleares. Sin embargo, en mediciones muy atentas de colisiones han venido apareciendo cada vez más resonancias moleculares de baja energía en las secciones eficaces. En efecto, Stephen Korotky y sus colaboradores, de Yale, Helmut Voit y sus colaboradores, de la Universidad de Erlangen, han determinado recientemente los números cuánticos o propiedades intrínsecas detalladas, de muchas de estas resonancias.

A las energías más altas aparecen abundantes picos estrechos en la sección eficaz de colisión ^{12}C-^{12}C. A partir del trabajo anterior de Torleif Ericsson, de la Organización Europea de Investigaciones Nucleares (CERN), y de Erich Vogt y sus colaboradores, de Chalk River, se sabía que muchos de estos picos podían ser productos de fluctuaciones estadísticas en los procesos nucleares que tenían lugar más que signos genuinos de moléculas nucleares. A altas energías, los núcleos en colisión pueden producir sistemas compuestos (fundidos) en uno cualquiera o en un gran número de estados de reacción compuestos sin que lleguen a formarse nunca moléculas nucleares. Estos estados compuestos tienen un amplio rango de vidas medias; por tanto, el principio de indeterminación de Heisenberg indica que los correspondientes picos resonantes varían de un modo notable; en particular, los más anchos de entre ellos se superponen. A una determinada energía para la cual la interacción sea genuinamente estadística, con la formación de un núcleo compuesto, un gran número de estos picos que se superponen sirven para definir la pauta de la sección eficaz. En circunstancias normales, los picos se su-

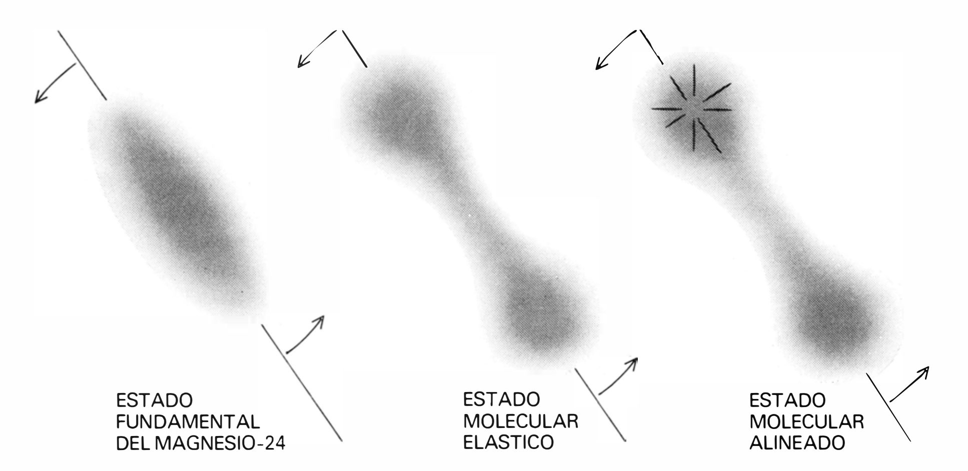


PRIMER ACELERADOR VAN DE GRAAFF, modelo "Emperor" en el A.W. Wright Nuclear Structure Laboratory de la Universidad de Yale, usado en el estudio de los fenómenos de moléculas nucleares. La fuente de iones que genera los proyectiles nucleares está al fondo, detrás de la gran vasija que sirven para acelerar los proyectiles. Se hace que los proyectiles nucleares adquieran una carga negativa añadiéndoles electrones. Como resultado, los proyectiles nucleares son atraídos hacia un terminal cargado positivamente en el centro de la vasija presurizada. Allí, los electrones son arrancados; los núcleos cargados positivamente resultantes son acelerados de nuevo por los efectos de repulsión electrostática hacia el extremo frontal de la vasija presurizada, que mide 30,48 m. (100 pies) de longitud y 6,09 m. de diámetro. (Fotografía publicada por cortesía de la Universidad de Yale.)

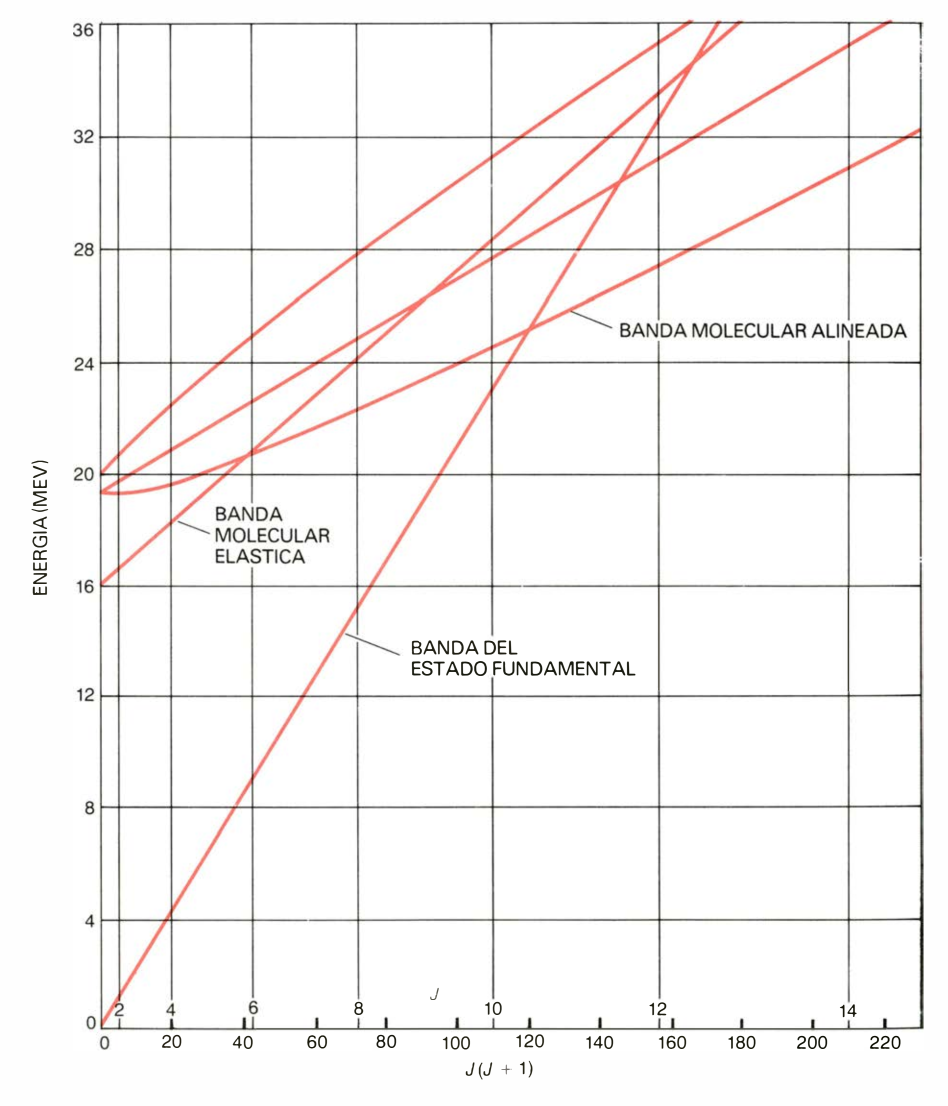
man al azar para producir una curva suave en la sección eficaz de la que sobresalen resonancias que representan configuraciones especiales, tales como las moléculas nucleares. Pero incluso en estas situaciones aleatorias pueden ocurrir a veces fluctuaciones estadísticas. Por ejemplo, cabe el que todos, salvo unos pocos, picos se cancelen mutuamente, dejando que las excepciones, o incluso un solo pico, determinen la sección eficaz. En algunos casos se han confundido tales picos con resonancias debidas a moléculas nucleares.

Buscando un camino que nos permitiera distinguir picos que fueran fenómenos genuinos de resonancia molecular de aquellos otros que fueran estadísticos, hubo que considerar en detalle modelos estadísticos de interacciones nucleares, que no se habían analizado antes en los estudios de las colisiones entre núcleos complejos. Aunque tales modelos, que habían sido desarrollados por Walter Hauser y Herman Feshbach, del Instituto de Tecnología de Massachusetts, se introdujeron originalmente para describir aspectos estadísticos de las reacciones nucleares inducidas por neutrones, nosotros descubrimos que podían reproducir, también con notable exactitud, los resultados de las reacciones inducidas por iones pesados, tales como los del carbono. En efecto, Daniel W. Shapiro y sus colaboradores en Yale demostraron que el modelo estadístico genera secciones eficaces tan ricas en picos como las secciones eficaces medidas a altas energías. Esto no significaba, sin embargo, que todos los picos de alta energía fuesen fluctuaciones estadísticas; sólo aclaraba que, en ausencia de mayores pruebas, ningún pico de alta energía podía identificarse como la señal de una molécula nuclear, cosa que se había hecho en algunas ocasiones.

Armados con los nuevos conocimientos de los aspectos estadísticos de las interacciones entre iones pesados, intentamos contestar la vieja pregunta de por qué los fenómenos relativos a las moléculas nucleares parecían necesitar de la participación del carbono. David R. Hanson y sus colaboradores en Yale habían demostrado que, de todas las combinaciones de núcleos que se fusionaban, la interacción de dos núcleos de carbono-12 era, con mucho, la que tenía el menor número de estados de reacción compuestos. Las colisiones entre núcleos de oxígeno-16, en cuya sección eficaz no aparecían resonancias, tenía 10.000 más estados de reacción compuestos que los núcleos podían ocupar. Sugerimos entonces que, aun suponiendo que existieran resonancias debi-



CONFIGURACIONES EN FORMA DE PESAS de gimnasio, representadas aquí de una forma esquemática (*centro* y *derecha*), que caracterizan las moléculas nucleares formadas en colisiones entre núcleos de carbono-12. En el estado molecular elástico ambos núcleos de carbono-12 se hallan en su estado fundamental (*centro*). En un estado molecular excitado típico, uno de los núcleos de carbono está en su primer estado excitado *(derecha)*. Incluso en el caso de que uno o ambos núcleos se hallen en un estado excitado, la molécula sigue manteniendo su forma. Tales moléculas nucleares difieren de la forma de un esferoide achatado que tiene el núcleo de magnesio-24 en el estado fundamental *(izquierda)*.



BANDAS ROTACIONALES, o conjuntos de estados, correspondientes a configuraciones nucleares particulares que giran en distintos estados cuantificados. Los estados que constituyen una banda son puntos que caen sobre una línea recta en una representación de la energía en función $J(J + 1)$, donde J es el número cuántico de momento angular total. La banda del estado fundamental, la banda molecular elástica y las bandas moleculares excitadas corresponden a las tres configuraciones de la figura superior. Las dos bandas más altas representan configuraciones en las que el spin del núcleo de carbono-12, en un primer estado excitado, es ya perpendicular o antiparalelo al momento angular del movimiento relativo de los dos núcleos. Se ha puesto la atención sobre el significado del cruce de la banda molecular elástica y de la banda molecular alineada, en las que los momentos angulares intrínseco y relativo son paralelos. Sólo se han probado experimentalmente algunos de los estados predichos.

das a moléculas nucleares en las colisiones de oxígeno-16, quedarían probablemente difuminadas y serían irreconocibles, debido a los numerosos estados competitivos, en las sección eficaz.

Esta hipótesis venía avalada por la observación de resonancias en las colisiones de núcleos de carbono-12 y oxígeno-16, que tenían también uno de los números más bajos de los estados de reacción compuestos. Confundían los resultados el hecho de que no aparecieran resonancias en las colisiones entre carbono-12 y berilio-9, y sólo muy débilmente en colisiones entre carbono-12 y carbono-13, a pesar de que ambos tenían un número bajo de estados de reacción compuestos. Creeemos ahora que, en estas dos clases de colisión, el comportamiento del neutrón de valencia, que está poco ligado, tanto en el carbono-13 como en el berilio-9, eclipsa y difumina los fenómenos de moléculas nucleares.

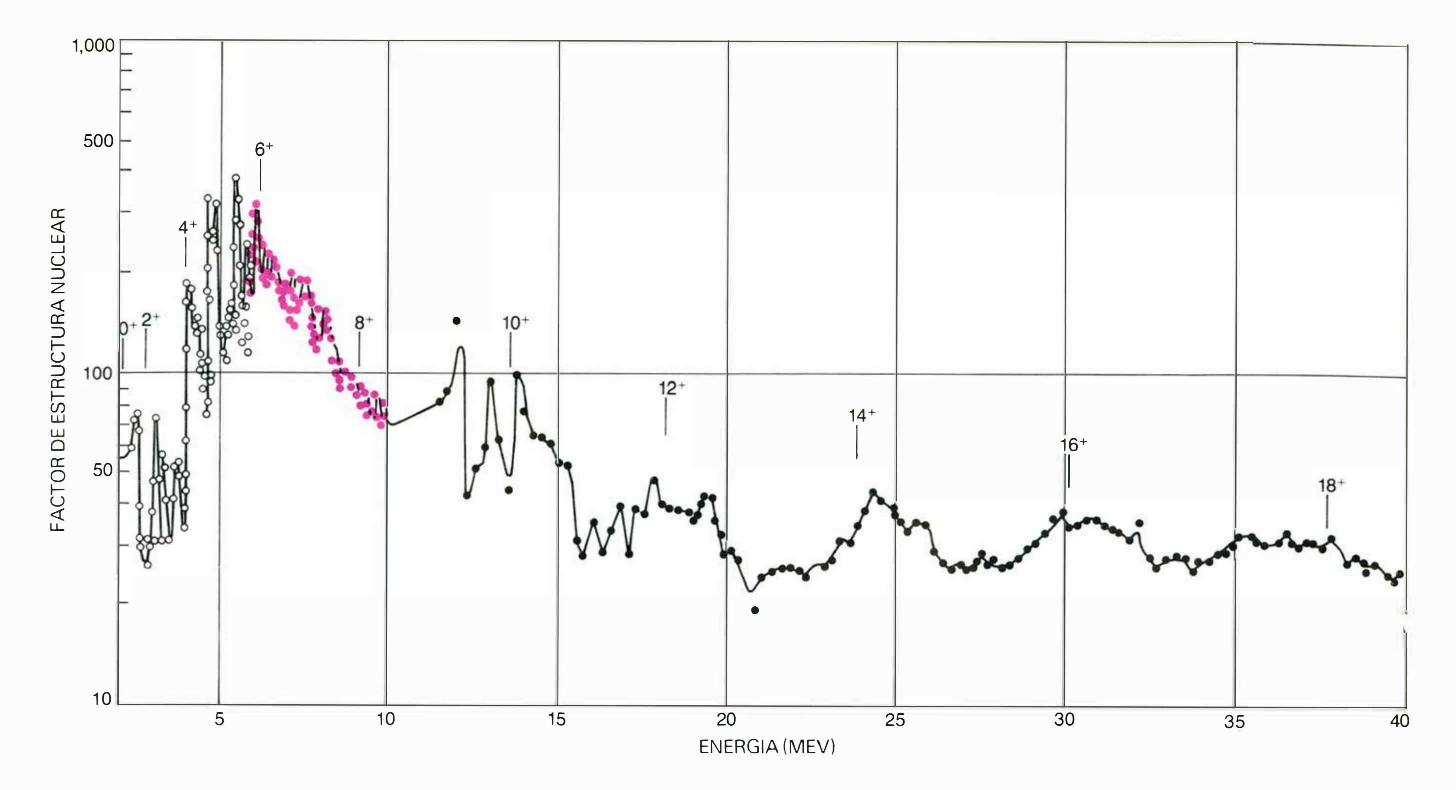
Al principio estábamos algo desanimados por nuestra hipótesis sobre las condiciones bajo las que se podrían observar moléculas nucleares, ya que predecía que dichas resonancias podían ser relativamente raras en física nuclear. Por fortuna, nuestra desazón no duró largo tiempo. Tan pronto como se desarrollaron nuevos métodos para obtener datos precisos sobre secciones eficaces, nosotros y otros encontramos fenómenos de resonancia molecular en muchas interacciones, incluyendo las del oxígeno-16 con oxígeno-16, carbono-12 con carbono-13, boro-11 con carbono-12, magnesio-24 con carbono-12, silicio-28 con oxígeno-16 y silicio-28 con silicio-28. Por supuesto, estas resonancias son generalmente menos acusadas que la de la interacción ^{12}C-^{12}C, donde los efectos competitivos resultan, con mucho, menores.

En los últimos años, los físicos nucleares de muchos laboratorios del mundo se han venido interesando por las moléculas nucleares. Este renacido interés obedece a desarrollos substancialmente nuevos, tanto experimentales como teóricos. Por la parte experimental, nuevos potentes ciclotrones y aceleradores electrostáticos han aumentado las energías a las que es posible reunir las partículas que chocan con alta precisión y han permitido medir secciones eficaces de colisión con elevada precisión y en muchos sistemas nuevos. Además, se han encontrado nuevos métodos para detectar productos nucleares especiales tales como el berilio-8, el cual, aunque tiene una vida media de sólo 10^{-16} segundos, suministra una prueba muy sensible de la estructura de las moléculas nucleares. Por lo que se refiere a la vertiente teórica, Yasuhisa Abe y sus colaboradores, de la Universidad de Kyoto, han encontrado que una versión más complicada del modelo de Imanishi predecía muchas más resonancias a baja energía correspondientes a moléculas nucleares, incluidas todas las conocidas por aquel entonces, que las que se suponía podía predecir el modelo original de Imanishi. Este conocimiento estimuló la búsqueda de aquellas resonancias predichas y que todavía no habían sido identificadas por vía experimental.

Desde que se descubrieron las moléculas nucleares, su momento angular fue una de las propiedades determinadas más importantes y menos ambiguas. De acuerdo con la teoría cuántica, el momento angular de cualquier sistema microscópico, como es el caso de una molécula nuclear, puede tener sólo ciertos valores discretos: $Jh/2\pi$, donde J es un entero y h la constante de Planck. Aunque la vida de una molécula nuclear es demasiado corta para poder medir su momento angular directamente, este momento puede determinarse por vía experimental a partir de la distribución angular de los núcleos que resultan de la fisión de la molécula nuclear o de la distribución angular de los productos nucleares en los que se desintegra la molécula nuclear. Si el modelo teórico de moléculas nucleares es válido, debería predecir correctamente los valores del momento angular determinados por vía experimental. Y lo ha hecho; ha sugerido, además, estados de resonancia adicionales, mucho de los cuales han sido ya detectados.

Walter Greiner y sus colaboradores, de la Universidad de Frankfurt, realizaron una contribución valiosa al hallar una nueva extensión del modelo original de Imanishi, que es particularmente importante a las energías más altas en que ahora se investiga. Encontraron que, a ciertas energías, pueden aparecer resonancias simultáneas en el movimiento relativo de los núcleos constituyentes y en el proceso de excitación mediante el que la energía se transmite del movimiento relativo a la excitación interna. Las resonancias simultáneas indican que, en estas condiciones, ha subido mucho la probabilidad de que se formen moléculas nucleares. Los datos existentes están de acuerdo con esta hipótesis.

COLISIONES RASANTES entre núcleos de carbono-12, que se sabe que tienen la forma de pomos de puerta, pueden hacer que cada uno de los núcleos gire alrededor de un eje perpendicular al eje de simetría del pomo. Se ha demostrado experimentalmente que la probabilidad de que ambos núcleos emerjan de la colisión de estados excitados de rotación es comparable a la probabilidad de que uno sólo de los núcleos sea puesto en tal estado. Ello sugiere una interacción tipo "rueda dentada", donde los dos núcleos de carbono se colocan en rotación igual y opuesta por fricción en el punto de contacto.



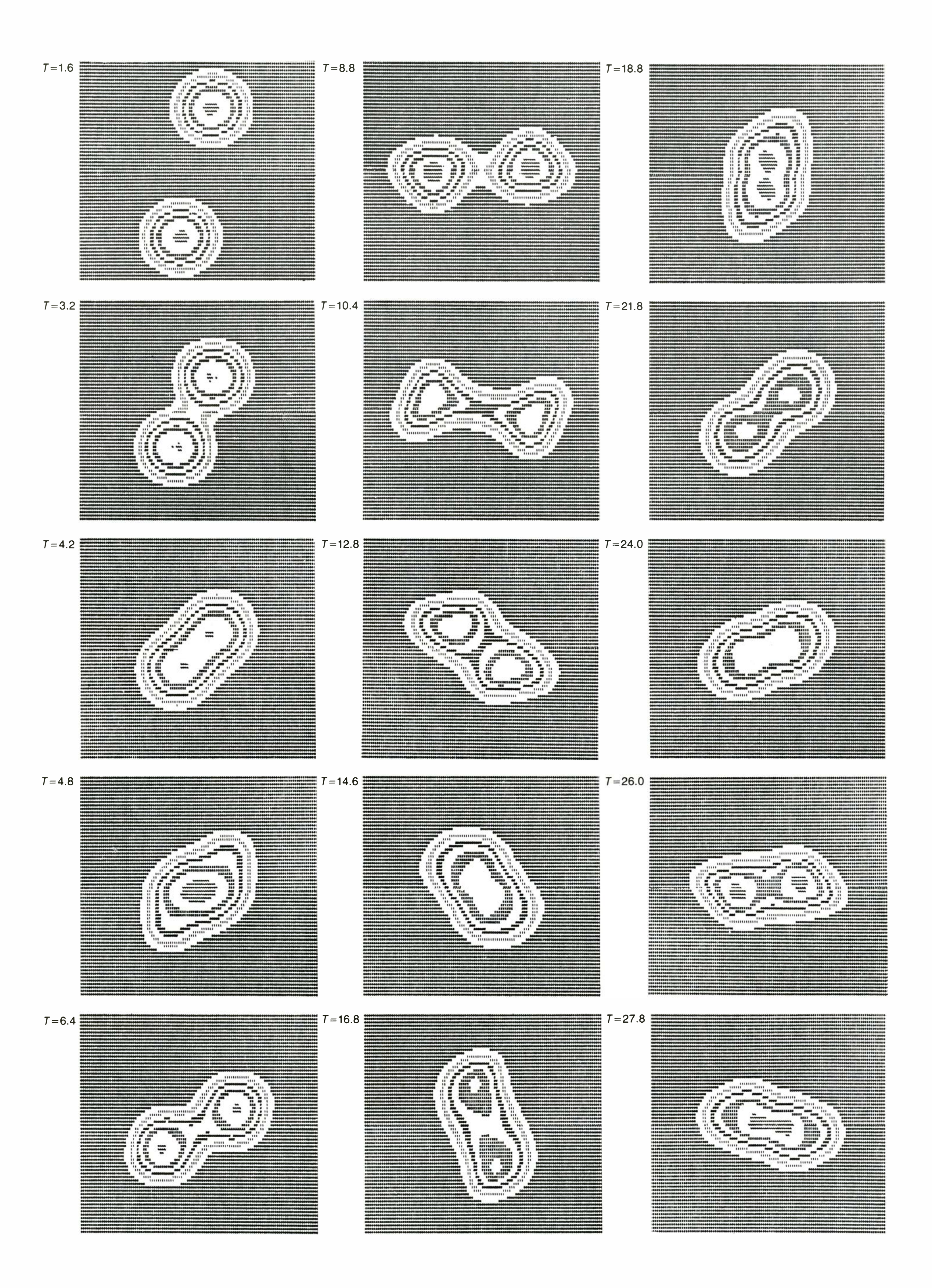
SECCION EFICAZ TOTAL de la interacción ^{12}C-^{12}C, representada en función de la energía disponible. Los círculos representan datos obtenidos en la Universidad de Münster, los puntos coloreados corresponden a datos proporcionados por Yale y los puntos negros son datos alcanzados en la Universidad del estado de Nueva York en Stony Brook. El factor de estructura nuclear representa aquella parte de la sección eficaz de interacción que depende sólo de la estructura de los núcleos de carbono. Las líneas verticales indican las energías a las que los modelos usuales han predicho resonancias moleculares elásticas. Cada resonancia se indica por el spin total en unidades $h/2\pi$, donde h es la constante de Planck. La aparente falta de subestructura en las resonancias de spin más elevado puede deberse a que las mediciones correspondientes se realizaron con precisión suficiente para revelar tal subestructura, si ésta se hallara presente. Las resonancias aparecen a lo largo del intervalo de energías de 40 MeV estudiado hasta hoy.

El trabajo investigador de numerosos físicos ha permitido esbozar una interpretación elegante de las propiedades rotacionales de la molécula nuclear ^{12}C-^{12}C. Tal molécula, en su configuración cuántica más simple, tiene la forma de unas pesas de gimnasia en rotación. Esta forma es muy distinta de la de un núcleo de magnesio-24 en su estado fundamental, o estado de menor energía; si bien el núcleo de magnesio-24 tiene el mismo número de protones y neutrones que dos núcleos de carbono-12, su forma semeja la de una pelota de rugby (un esferoide achatado). Consecuencia general de la mecánica cuántica es que una entidad no esférica con la forma de unas pesas de gimnasia o una pelota de rugby pueda conservar su estructura interna al aumentar su velocidad de rotación, pero únicamente para valores de la velocidad angular discretos y cuantificados. Una serie de tales estados cuantificados constituye una banda rotacional, bien conocida en física de moléculas atómicas. Los estados con una velocidad angular dada y, por tanto, con un momento angular determinado, tienen energías que son proporcionales a $J(J+1)$, donde J simboliza el número cuántico de momento angular. Las moléculas nucleares que pertenecen a una cierta banda ostentan todas la misma estructura interna; y difieren unas de otras sólo en lo referente a la velocidad de rotación de dicha estructura.

Puesto que las energías de los estados de las moléculas nucleares son proporcionales a $J(J+1)$, los miembros de una banda rotacional caerán sobre una línea recta en una representación gráfica de la energía en función de $J(J+1)$. En la banda molecular elástica, o menos energética, los núcleos de carbono que forman la molécula no están excitados; es decir, también ellos se encuentran en el estado menos energético. Las bandas moleculares excitadas, donde uno o ambos carbonos están también en estados excitados, tienen una estructura más compleja que revela las distintas formas en que el momento angular de las pesas de gimnasia puede acoplarse con el momento angular de los núcleos constituyentes y excitados para definir el momento angular de la molécula nuclear como un todo.

De esta interpretación sorprende el que una de las bandas moleculares excitadas, una banda molecular alineada en la que las pesas de gimnasia ruedan alrededor de un eje que está alineado con el spin intrínseco de los núcleos excitados constituyentes, cruza la banda molecular elástica en un punto. Más aún, bandas moleculares alineadas, que corresponden a excitaciones intrínsecas o excitaciones mutuas de los núcleos de mayor energía, cruzan la banda molecular elástica a valores, todavía más elevados, de los momentos angulares y de la energía. ¿Cuál es el significado de estos puntos de cruce? A los momentos angulares y energías a los que se cruzan las dos bandas, la molécula toma una configuración que es una mezcla de las características de las dos bandas. Por ejemplo, en el punto donde las bandas molecular elástica y la alineada se cruzan, la molécula está parte del tiempo en cada uno de estos estados. En este punto, una molécula elástica, formada en la colisión entre dos núcleos de carbono no excitados, podría convertirse en una configuración alineada sin gastar energía o momento angular.

En tales circunstancias ha aumentado grandemente la probabilidad de que la molécula alineada se disocie en dos núcleos de carbono, uno de los cuales se encontraría en el estado excitado que caracteriza la banda que cruza. Esta interpretación de bandas predice que, cerca de los cruces de bandas, las resonancias moleculares se dividirán en dos picos separados, un pico para cada banda. Tales picos dobles han sido observados recientemente en moléculas nucleares ^{12}C-^{12}C



T=1.6
T=8.8
T=18.8
T=3.2
T=10.4
T=21.8
T=4.2
T=12.8
T=24.0
T=4.8
T=14.6
T=26.0
T=6.4
T=16.8
T=27.8

y ^{12}C-^{16}O a energías adecuadas, lo que dio más fuerza a la credibilidad de la interpretación de bandas moleculares y estimuló una amplia búsqueda de los desdoblamientos predichos en otras moléculas nucleares mucho más excitadas.

La extensión del modelo de Imanishi, que tiene en cuenta el acoplamiento entre el movimiento de las pesas de gimnasia y los movimientos intrínsecos de los núcleos contituyentes, así como el intercambio de energía entre estas dos clases de movimiento, ha dado origen a predicciones que han sido confirmadas en muchos experimentos. Abe y sus colaboradores usaron este modelo para predecir un triplete entonces no detectado de estados resonantes a energías exactamente por encima de la barrera de Coulomb, que es la energía necesaria para contrarrestar exactamente la repulsión electrostática entre los núcleos. Más tarde, se halló el triplete. Aunque no se han encontrado cierto número de resonancias predichas por el modelo, el acuerdo general obtenido hasta ahora entre las predicciones y los datos medidos resulta esperanzador.

Si las resonancias observadas se corresponden o no, de un modo preciso, con la localización de las originalmente predichas no constituye ya ninguna cuestión importante, por la razón de que unos ajustes, relativamente pequeños, en los parámetros del modelo permitirán que la interpretación de bandas acople los datos. El modelo también exige que cuando una molécula nuclear de una configuración alineada se disocie a través de colisiones inelásticas, los núcleos de carbono excitado que salgan girarán todos en una dirección común. Se está trabajando en las difíciles mediciones que se precisan para confirmar este fenómeno; los resultados habrán de ayudar a establecer la validez física de las bandas.

La interpretación de las bandas, la descripción de resonancias simultáneas y la gran cantidad de datos nuevos que las avalan han resuelto algunos de los aspectos paradójicos del modelo original de Imanishi y los datos de los primeros tiempos. De los datos primitivos se sugirió, por extrapolación, que en el centro de ciertas estrellas se quemaba el carbono a un ritmo que superaba en más de 10 veces el predicho por cualquier modelo de evolución estelar; los trabajos recientes indican, sin embargo, que las difíciles mediciones primitivas contenían errores experimentales y que la extrapolación era incorrecta. El ritmo a que se quema el carbono, tal como lo postulan los datos actuales, es del mismo orden de magnitud que los ritmos propuestos por estos modelos astrofísicos. El modelo original de Imanishi había avanzado la hipótesis, poco física en apariencia, de que las superficies nucleares que intervenían en los fenómenos de moléculas nucleares penetraban frecuentemente unas en otras sin perturbarse; gracias a trabajos recientes, tal hipótesis resulta más comprensible. Esas investigaciones manifiestan que cuando se toman en cuenta de forma apropiada los efectos debidos al momento angular, las superficies nucleares se muestran realmente muy transparentes.

Una parte destacada del trabajo más reciente en este campo se centra en la medición de las secciones eficaces de fusión total: las probabilidades de que los núcleos que intervienen en las colisiones se unan para formar un único compuesto. Puesto que la atracción nuclear entre los dos núcleos es muy fuerte a las distancias de las moléculas nucleares, se esperaba que la formación de moléculas nucleares incrementara la probabilidad de fusión de los núcleos, como en el caso de dos núcleos de carbono que se unieran para formar un núcleo compuesto de magnesio. Esta predicción teórica se ha confirmado ya, por vía experimental, en las colisiones ^{12}C-^{12}C y ^{12}C-^{16}O; en ellas, las resonancias de fusión parecen estar en acuerdo satisfactorio con los estados que predice la interpretación de bandas moleculares.

En los estudios de primera hora sobre la colisión ^{16}O-^{16}O no se observaron picos que sugirieran la formación de moléculas nucleares; probablemente ello obedecía a que tales picos habían quedado difuminados por el gran número de efectos de núcleo compuesto que intervenían. Ahora, sin embargo, han aparecido estructuras resonantes indicativas de las moléculas nucleares en los datos de fusión de ^{16}O-^{16}O obtenidos por Peter Sperr y sus colaboradores, en el Argonne National Laboratory, y por Florent Haas y sus colaboradores, en la Universidad de Estrasburgo. Pero no todos los datos de fusión muestran estructuras resonantes. En colisiones entre núcleos de oxígeno 18 (^{18}O), por ejemplo, la presencia de dos neutrones de valencia fuera del núcleo básico de oxígeno-16 difuminan resonancias que, de otro modo, podrían revelarse. La presencia de dos neutrones de valencia y de un protón de valencia en el flúor-19 (^{19}F) puede explicar por qué no hay estructura resonante en las colisiones ^{19}F-^{12}C, pero aquí la explicación se hace menos convincente. Las difíciles mediciones de las secciones eficaces de estas interacciones se hallan en sus primeros pasos, y quizá lleguen a detectarse estructuras resonantes.

Recientemente, nuevas mediciones han revelado la existencia de picos irregulares en las secciones eficaces de la reacción ^{12}C-^{12}C a altas energías. Cada pico irregular parece ser la envolvente de un gran número de resonancias agudas y cercanas, aunque espaciadas de forma irregular, que pueden corresponder a estados específicos del núcleo compuesto de magnesio-24 (^{24}Mg). Feshbach ha sugerido que estos picos irregulares podrían ser muy bien una ilustración clásica de un modelo de "portal" de interacciones nucleares, que él desarrolló para describir las interacciones que comprometían a núcleos mucho más ligeros que los de carbono. En el modelo de portal se supone que el carbono-12 blanco y el carbono-12 proyectil se combinan inicialmente para formar un estado de particular simplicidad, que o bien se disocia de nuevo en el blanco y proyectil originales, o bien atraviesa un portal figurado y se desarrolla paso a paso a través de una serie de configuraciones más complejas hacia un equilibrio estadístico que es característico de un núcleo compuesto de magnesio muy excitado.

Queda mucho trabajo por realizar para substanciar el modelo de portal de interacciones de iones pesados. Así, de acuerdo con la teoría cuántica, los estados compuestos subyacentes de resonancia bien definidos en los que el estado portal puede disolverse deben tener todos el mismo spin y paridad (simetría intrínseca) que el estado molecular portal, pues se conservan el spin total y la paridad. Si los valores del spin y de la paridad determinados experimentalmente satisfacen esta exigencia, el modelo portal recibirá un fuerte espaldarazo. Además, para identificar de forma concluyente un estado portal como estado molecular hay que probar experimentalmente que cada uno de los estados bien definidos tiene una fracción substancial de la configuración molecu-

LAS LINEAS DE CONTORNO indican la densidad calculada en la dirección perpendicular al plano de colisión para choques entre núcleos de oxígeno-16 con una energía de 52,5 MeV y con un momento angular de $13h/2\pi$. La unidad de tiempo es 10^{-22} segundos. Estos contornos, calculados por Hubert C. Flocard, Steven Koonin y Morton Weiss, del Lawrence Livermore Laboratory de la Universidad de California, ilustran cómo se espera que los núcleos de oxígeno-16 se junten para formar una molécula, que gira y vibra después de los 3×10^{-21} segundos que se muestran en la figura de la página opuesta.

lar en su estructura intrínseca. Se están llevando a cabo mediciones a este respecto.

Resumamos lo que los experimentos han revelado acerca de la estructura y la dinámica de las moléculas nucleares ^{12}C-^{12}C, por tratarse de las clases de moléculas nucleares más estudiadas. Han aparecido resonancias moleculares cada vez que se ha escudriñado detenidamente un nuevo intervalo de energías. A más altas energías, muy por encima de la barrera de Coulomb, las resonancias moleculares aparecen como picos fragmentados con una anchura de unos pocos millones de electronvolt, que están compuestos de subresonancias estrechas cuya anchura típica alcanza algunos cientos de miles de electronvolt. Gran parte del trabajo a propósito de las moléculas de carbono se ha centrado en lo que sucede cuando uno o los dos núcleos que colisionan se excitan hasta el estado excitado más bajo del núcleo de carbono a 4,43 MeV. Sin embargo, cuando la energía de los núcleos que chocan se aumenta hacia el máximo que se puede alcanzar con los actuales aceleradores de alta precisión, se descubre que intervienen una gran variedad de estados excitados. Así, Karl A. Erb y sus colaboradores, de Yale, y Ken Kantori y sus colaboradores han visto un núcleo de carbono absorber 14,05 MeV y moverse a un estado de spin más alto, mientras el otro núcleo de carbono permanecía en el estado fundamental o absorbía 4,43 MeV. En el último caso, se transfirieron 18,48 MeV de energía, que es una cantidad extraordinariamente grande para fenómenos nucleares, desde el movimiento relativo hasta la excitación interna mediante el simple mecanismo de girar ambos núcleos de carbono a estados de spin más altos.

Las colisiones en las que se transportan a la vez a ambos núcleos a estados rotacionales más altos tienen sorprendentemente largas secciones eficaces; la probabilidad de que ambos núcleos sean excitados es comparable a la probabilidad de que uno sólo de los núcleos lo sea. Esto contrasta de un modo radical con lo que ocurre en las habituales colisiones nucleares; para dar cuenta de esta anomalía hemos propuesto una interacción de "rueda dentada", en la que los dos núcleos se colocan en rotación igual y opuesta debido a la fricción en el punto de contacto en las colisiones fuera del centro. No disponemos aún de pruebas experimentales que muestren que los momentos angulares en estas colisiones están alineados en sentidos opuestos, y quizás aportarían un argumento concluyente de este simple mecanismo de excitación.

El trabajo teórico con moléculas nucleares se ha concentrado, sobre todo, en las propiedades macroscópicas, tales como las características rotacionales. Desde hace poco tiempo se han podido investigar también las propiedades microscópicas, como el distinto comportamiento de los protones y neutrones dentro de los núcleos que chocan. Así, en el Oak Ridge National Laboratory, Joachim Maruhn y Ronald Y. Cusson han calculado hasta qué punto cambia la densidad de materia en una colisión frontal ^{12}C-^{12}C durante el breve tiempo que dura la interacción. Los fenómenos que intervienen son tan complejos que Maruhn y Cusson han tenido que recurrir a ciertas aproximaciones simplificadoras en sus cálculos. En este sentido, se han limitado a colisiones frontales para no tener que considerar los efectos del momento angular, y trataron cada protón y neutrón como si "sintieran" únicamente la fuerza media de los 23 protones y neutrones restantes. En particular, ignoraron todos los efectos que podrían comportar los choques individuales entre los protones y neutrones en las configuraciones orbitales dentro de los núcleos implicados. Resultó especialmente satisfactorio que surgieran configuraciones de moléculas nucleares incluso a partir de estos toscos cálculos microscópicos; no había ninguna razón a priori para esperar que tal ocurriera.

En el Instituto de Tecnología de Massachusetts, Yale, Oak Ridge, Universidad de California en Berkeley y en el Lawrence Livermore Laboratory se están realizando trabajos que prometen aumentar la precisión de estos cálculos. Los cálculos primitivos se aplicaron a colisiones en las que los núcleos de carbono se fusionaban en un núcleo muy excitado de magnesio-24, vibraba colectivamente y después se separaba de nuevo en dos núcleos de carbono. Como la colisión se suponía frontal, los cálculos no revelan nada sobre la rotación del sistema. Después de la colisión la densidad de cada núcleo varía de un punto a otro, de una forma irregular si la comparamos con la variación suave de la densidad antes de la colisión. La densidad irregular significa que la energía se ha transferido del movimiento relativo a complicadas excitaciones internas. Hubert C. Flocard, Steven Koonin y Morton Weiss, de Livermore, han llevado a cabo cálculos de la densidad más ambiciosos para una colisión no frontal entre núcleos de oxígeno, y, por tanto, incluía efectos rotacionales y vibracionales.

Los cálculos sugieren que, a energías relativamente bajas, la densidad de los núcleos que chocan nunca se eleva mucho por encima del valor de la densidad en el centro de un solo núcleo de carbono. Tanto a bajas como a altas energías, se forma un largo cuello durante la colisión para ligar los núcleos de carbono por un tiempo relativamente largo en una configuración molecular que gira y vibra. Los cálculos se han limitado en general a los primeros estadios de la colisión, debido a la impotencia de los ordenadores existentes que obligan a cortar los cálculos mucho antes de que la colisión haya llegado a sus fases finales.

Urich Mosel y sus colaboradores, de la Universidad Justus Liebig, de Giessen, han trabajado con un modelo de capas para describir el núcleo. Los protones y neutrones del núcleo se suponen distribuidos en capas análogas a las capas electrónicas de los átomos, al objeto de seguir los movimientos individuales de los neutrones y protones durante la colisión de núcleos de carbono. Han obtenido así resultados parecidos a los de Livermore. Aunque las técnicas para hacer estos cálculos microscópicos están aún en sus balbuceos, prometen, con el tiempo, lograr una imagen detallada del distinto papel que los neutrones y protones desempeñan en las colisiones de iones pesados. Tal imagen constituiría un gran avance en el camino de la comprensión de la dinámica y la estructura nuclear.

Desde lo que empezó siendo una simple curiosidad en 1960, las moléculas nucleares se han revelado como un aspecto común de la estructura y dinámica nuclear de los sistemas de iones pesados estudiados hasta el presente. Las moléculas nucleares han aportado una respuesta joven a una vieja cuestión: ¿Qué le sucede a un núcleo cuando se le suministra energía? No hace muchos años, Peter D. Parker y sus colaboradores de Yale demostraron que, en una primera aproximación, los espectros de estados excitados de núcleos ligeros podían interpretarse como si el radio del núcleo aumentara alrededor de 10^{-14} centímetros por cada millón de electronvolt de energía que se añadiera. En otras palabras, el núcleo se expansiona al calentarse. Pero si esto fuera lo único que sucede, el núcleo resultaría un objeto realmente aburrido. Se ha encontrado ahora, sin embargo, que a ciertas energías específicas, muy por encima de las necesarias para que se desprenda un neutrón, los protones y neutrones de un núcleo forman agregados y configuraciones moleculares en las que subgrupos se unen para formar núcleos bien definidos que se mueven en un sencillo minué unos respecto a otros. Esto abre un área totalmente nueva de física nuclear.

Centrales energéticas de células de combustible

Estas células generan electricidad directamente de un combustible sin emitir contaminantes. Se está potenciando su escala al objeto de determinar si pueden resultar rentables para el suministro de fluido

Arnold P. Fickett

En la próxima década el suministro eléctrico de los Estados Unidos requerirá generadores de energía que cumplan ciertas exigencias no usuales: elevada eficacia, baja emisión de contaminantes, operación silenciosa e instalación rápida. Los generadores deberán suministrar electricidad en áreas urbanas donde los generadores convencionales resultarían inaceptables por razones de índole ambiental. En la célula de combustible encontramos un candidato idóneo para ello.

La célula de combustible convierte la energía química de un combustible en electricidad directamente, sin ningún ciclo de combustión intermedio. Para ser un aparato inventado en 1839 (por Sir William Grove, un jurista británico que hizo notables contribuciones a la ciencia) ha transcurrido mucho tiempo hasta que volviera al primer plano de la actualidad. La célula de combustible suministró la energía para los ingenios espaciales Géminis y Apolo, pero ésta ha sido una aplicación exótica y cara. Ahora, la célula de combustible, notablemente mejorada con respecto a las versiones que se utilizaron en las aeronaves, y formando conjuntos pluricelulares 1000 veces mayores, parece haber alcanzado una etapa donde puede contribuir de una manera destacada al suministro de electricidad de una nación. En efecto, la célula de combustible se contempla en la industria de las fuerzas eléctricas de los Estados Unidos como una alternativa importante para atender la demanda creciente esperada para las dos décadas próximas; con la ventaja de que no altera lesivamente el medio ambiente y ahorra combustible.

A corto plazo se espera que los sistemas de células de combustible puedan desarrollarse en los años 80 para atender cargas de horas punta y para lo que la industria de suministro denomina "siguiendo la carga". Los generadores de horas punta se ponen en marcha sólo cuando se incrementa por un tiempo dado la necesidad de energía adicional. Los generadores siguiendo la carga se ponen en marcha diariamente y trabajan la mayor parte del tiempo acoplados con las oscilaciones diurnas de la carga; se pueden parar de noche y durante el fin de semana. Como generadores de horas punta temporales o siguiendo la carga, las centrales energéticas de células de combustible utilizarían más eficazmente los combustibles líquidos procedentes del petróleo o carbón, aprovechando mejor el suministro de combustible fósil de una nación. A largo plazo, las células de combustible integradas con gasificadores de carbón aparecen como posibles grandes centrales que suministren la carga base, centrales que operarían las veinticuatro horas y utilizarían carbón directamente.

Una célula de combustible consta de dos electrodos –un electrodo positivo, el cátodo, y un electrodo negativo, el ánodo– separados por un electrolito, que transmite iones pero no electrones. Un combustible, típicamente el hidrógeno, se suministra al ánodo y el oxígeno (del aire) se suministra al cátodo.

Un catalizador en el ánodo poroso hace que las moléculas de hidrógeno (H_2) se disocien en iones hidrógeno (H^+) y electrones. Si el electrolito es ácido, los iones hidrógeno emigran a través del mismo hacia el cátodo. donde reaccionan con los electrones (suministrados a través de la carga del circuito externo) y oxígeno para formar agua (H_2O). La naturaleza del ion emigrante depende del electrolito. Los iones hidrógeno emigran del ánodo al cátodo en un electrolito ácido; los iones hidroxilo (OH^-) emigran del cátodo al ánodo en un electrolito alcalino; los iones carbonato (CO_3^{--}) emigran del cátodo al ánodo en un electrolito de sales de carbonatos y los iones oxígeno (O^{--}) emigran del cátodo al ánodo en un electrolito de óxido sólido. En cada caso, las reacciones producen electrones que, si los electrodos están conectados por un conductor, circulan de un electrodo al otro a través del circuito externo. Puesto que un flujo de electrones constituye una corriente eléctrica, los electrones que se mueven en el circuito pueden utilizarse para alumbrar una lámpara, accionar un motor eléctrico, etcétera.

El voltaje máximo de corriente continua producido por una célula de combustible es una función termodinámica del combustible y del oxidante. Para una célula que trabaje con hidrógeno y oxígeno, el voltaje teórico a presión y temperatura ordinarias es 1,23 volt. El voltaje real será de 0,6 a 0,85 volt a causa de las pérdidas producidas en el interior de la célula. La corriente producida está controlada por la velocidad de las reacciones electroquímicas (la oxidación del hidrógeno, que implica la separación de electrones de los átomos de hidrógeno, y la reducción del oxígeno, que implica la adición de electrones a los átomos de oxígeno) y también por el área superficial disponible para las reacciones.

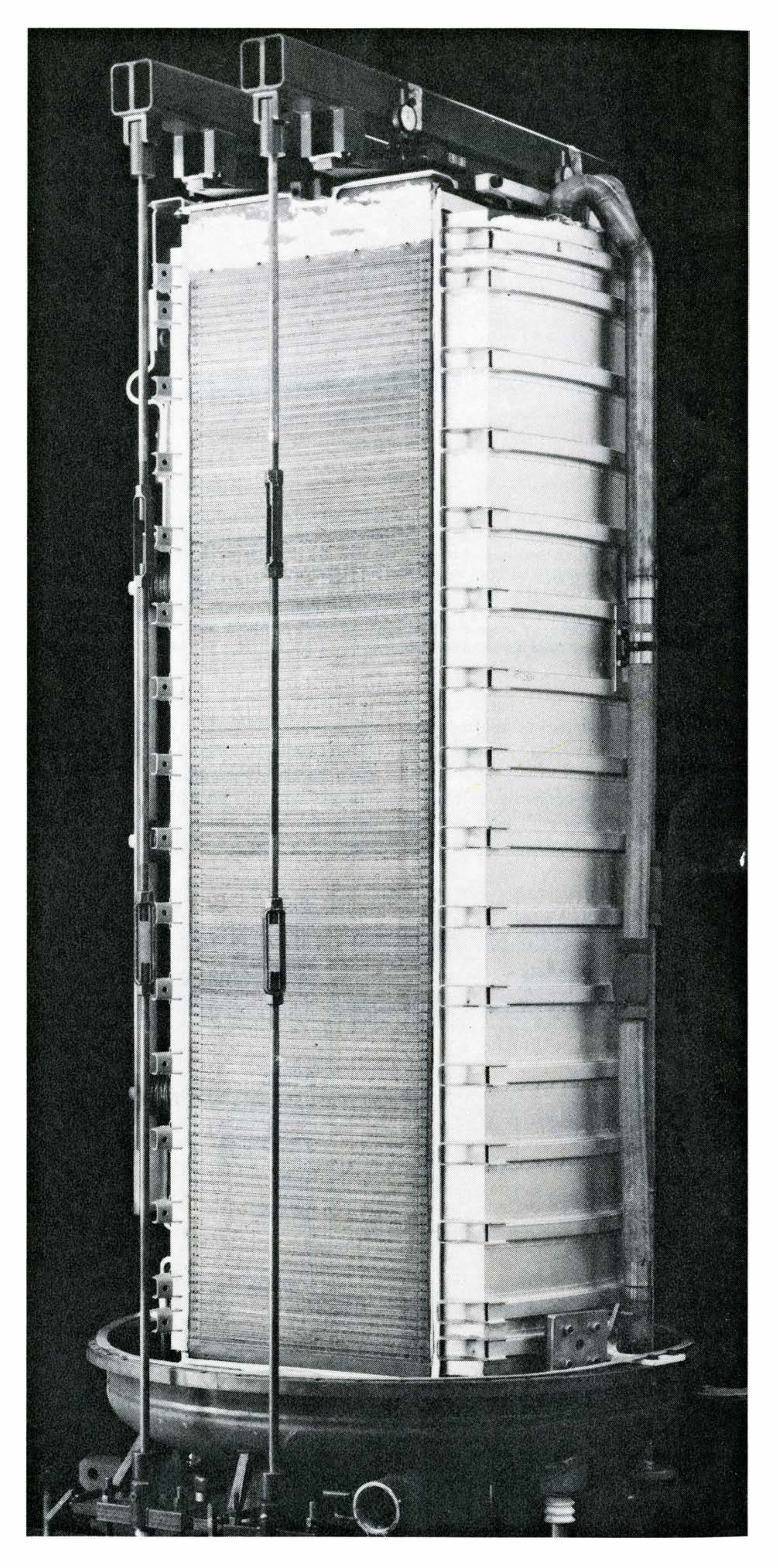
Puede fabricarse una célula sencilla de combustible colocando dos barras de carbón catalizadas con pequeñas cantidades de platino en un vaso que contenga ácido sulfúrico. Si se burbujea hidrógeno a través de la superficie de una de las barras y oxígeno a través de la superficie de la otra, se generará co-

rriente continua de aproximadamente un volt como consecuencia de la reacción electroquímica entre el hidrógeno y el oxígeno. El ácido sulfúrico opera de electrolito. Se puede observar el voltaje conectando un voltímetro entre las barras de carbón. Si las barras están conectadas a través de un circuito externo, los electrones fluirán a través del mismo, generando energía eléctrica.

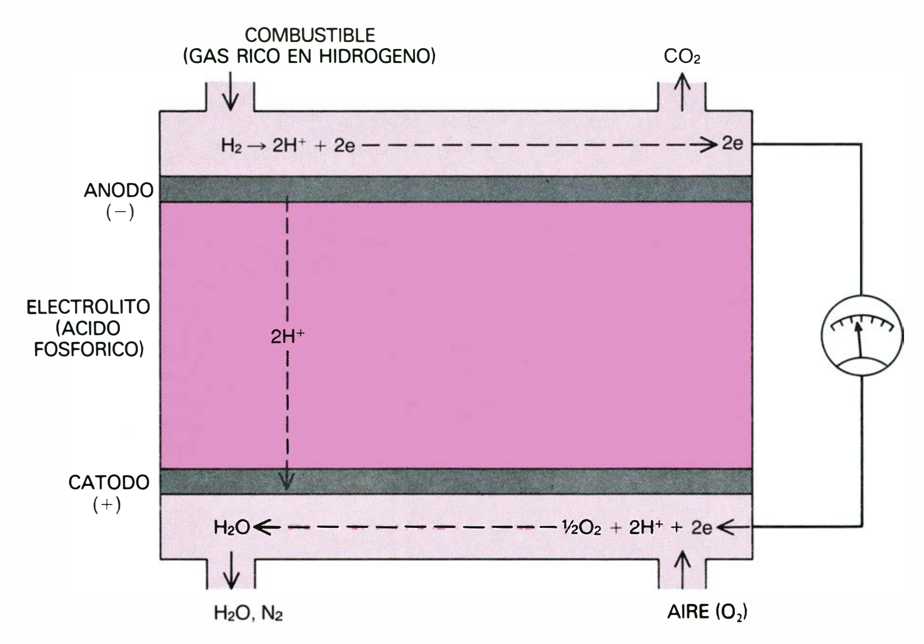
Aunque esta célula de combustión producirá un voltaje, no resulta práctica por varias razones. Las barras de carbón, al funcionar como electrodos, no ofrecen una gran área superficial para las reacciones. Además, el ácido sulfúrico no es ningún electrolito estable, y los vasos no son fáciles de unir en grandes grupos multicelulares.

Puede obtenerse una configuración práctica para una célula de combustible colocando el electrolito en una matriz o papel secante delgado situado entre dos electrodos porosos. (La porosidad da a los electrodos una gran área superficial.) Los catalizadores para las respectivas reacciones electroquímicas se incorporan en los electrodos. Las distintas células de combustible se combinan para formar un "conjunto" o paquete pluricelular cuyo voltaje de salida equivale al producto del voltaje de una célula por el número de las mismas. Una "sección energética", reunión de células de combustible diseñada para obtener una cantidad importante de energía, puede constar de uno o más conjuntos.

Si una central de energía, construida a base de células de combustible, ha de ser un generador útil de potencia para un suministro eléctrico, debe utilizar el tipo de combustible que sea económicamente aprovechable para el suministro y producir corriente alterna compatible con el sistema de transmisión del suministro. Por tanto, una central energética de células de combustible tendrá que comprender no sólo células de combustible sino también un tratamiento del combustible y un acondicionador de la energía. El tratamiento del combustible convierte un combustible de suministro, el



"CONJUNTO" DE CELULAS de combustible fotografiado en la planta del fabricante, la United Technologies Corporation. Cada capa del conjunto es una célula de combustible capaz de producir 500 watt de potencia a 0,65 volt. La producción del conjunto es dicha potencia (o voltaje) tantas veces como número de células hay en el conjunto (aproximadamente 500). Este conjunto es uno de los 20 que se instalarán en la planta de 4,5 megawatt, que se construirá en la ciudad de Nueva York para un ensayo avanzado de la tecnología. Las células tienen ácido fosfórico como electrolito y utilizarán como combustible un gas rico en hidrógeno que se obtiene a partir de nafta.

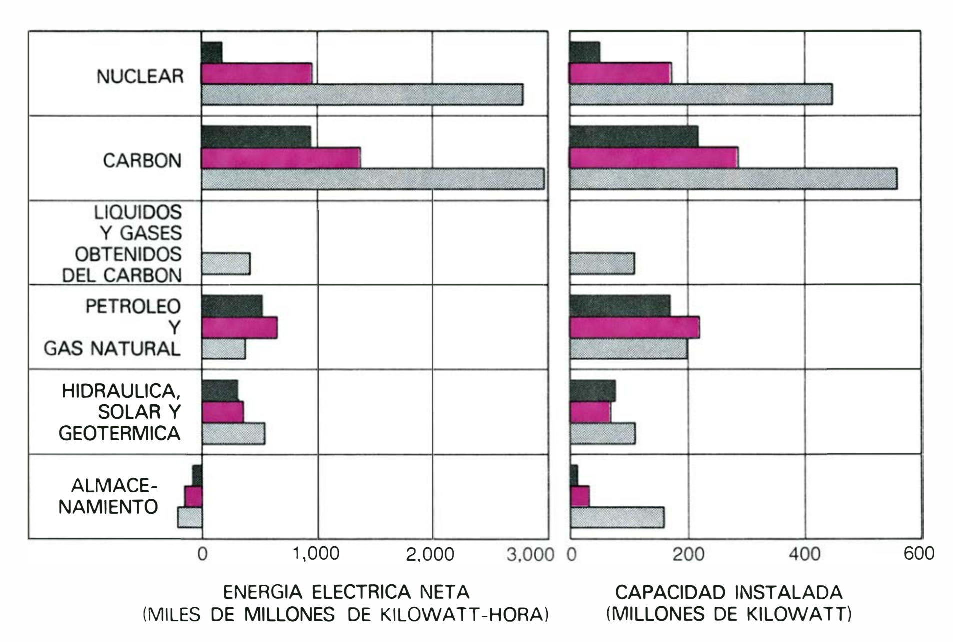


FUNCIONAMIENTO DE UNA CELULA DE COMBUSTIBLE: está basado en las reacciones electroquímicas entre un combustible, en este caso un gas rico en hidrógeno obtenido a partir de carbón o petróleo, y un oxidante, en este caso el oxígeno del aire. En el ánodo de la célula las moléculas de hidrógeno se oxidan (se separa un electrón de cada uno de los dos átomos de hidrógeno de la molécula) para formar iones hidrógeno. Los iones se transportan a través del electrolito hasta el cátodo, y los electrones circulan a través de un circuito externo hasta el cátodo, produciendo energía. En el cátodo los iones hidrógeno, electrones y oxígeno forman agua. La célula convierte, pues, la energía química de un combustible en electricidad directamente, sin recurrir a ningún ciclo de combustión intermedio.

carbón por ejemplo, en un gas rico en hidrógeno, y el acondicionador de energía convierte la corriente continua en corriente alterna.

El coste y la eficacia de la central energética en su totalidad dependen, en gran parte, del rendimiento de cada célula individual. Para deducir una estimación aproximada de la eficacia global de una central energética, desde el combustible hasta la energía en forma de corriente alterna, podemos recurrir a la ecuación $N_p = 59\ V_c$, donde N_p es la eficacia de la planta (en porcentaje) y V_c, es el voltaje de una célula. Esta relación es precisa, con un 5 por ciento de margen, e ilustra sobre la importancia del voltaje de una sola célula en el establecimiento de la eficacia global de una planta.

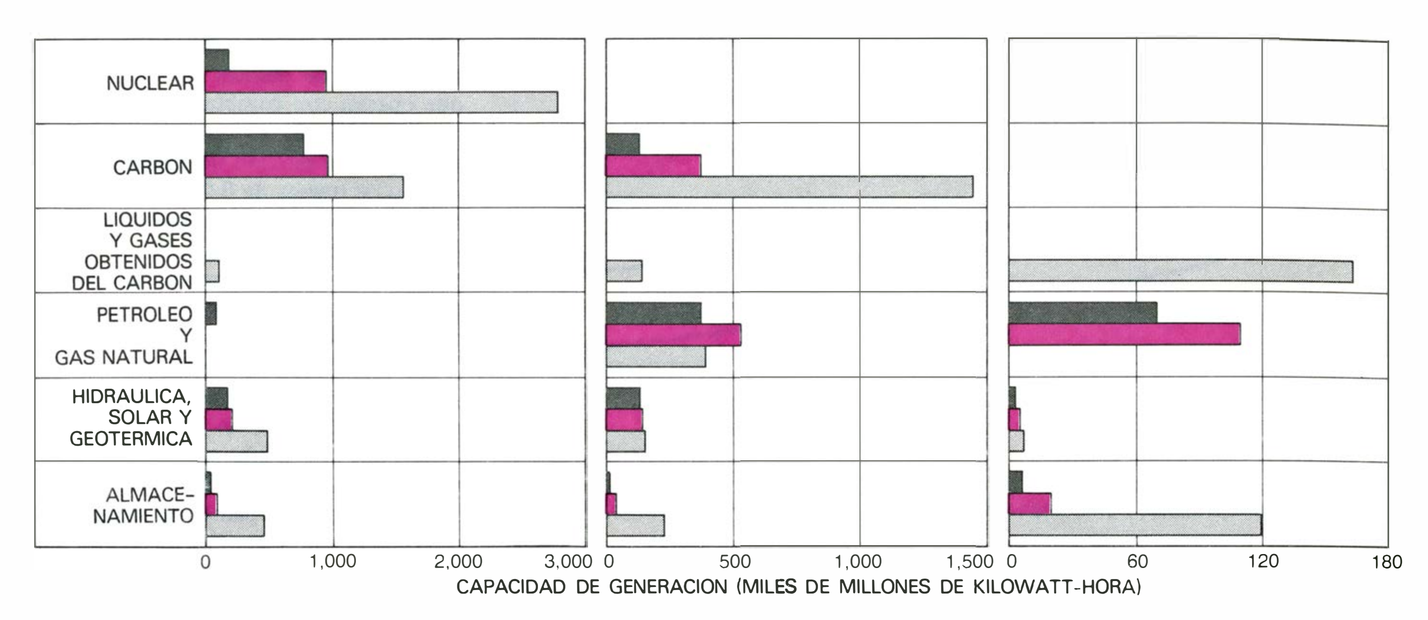
El coste de capital de una sección energética, expresada en dólares por kilovatt de capacidad, es inversamente proporcional a la densidad energética de una sola célula. Por tanto, la aportación energética de la célula establece de hecho el coste de capital de la sección energética. Típicamente la sección energética costará 100 dólares por kilovatt, lo que representa de 12 a un 30 por ciento del coste de la central energética, dependiendo de la complejidad del sistema de tratamiento del combustible. El procesado de carbón es mucho más complejo que el procesado de petróleo; por tanto, una central energética que utilice carbón tendrá un coste total superior que otra que utilice petróleo. Los objetivos de la mayoría de programas sobre células de combustible son mejorar el rendimiento de la célula, reducir su coste por unidad de área y aumentar su seguridad y resistencia.



DEMANDA Y DISPONIBILIDAD DE ELECTRICIDAD correspondientes a 1976 *(gris obscuro)* y, según previsiones del Electric Power Research Institute, a 1985 *(color)* y al año 2000 *(gris claro)*. Los cálculos suponen una reducción del 17 por ciento en la demanda total merced a la conservación. Las proyecciones son para las plantas nucleares y de carbón que suministrarán el 82 por ciento de la electricidad producida. Los fallos en el suministro de electricidad producida por las plantas nucleares o de carbón resultantes de problemas de localización o ambientales habrán de resolverse mediante la producción de otras formas de fuentes de energía tales como líquidos y gases. Las células de combustible, inventadas en 1839 por Sir William Grove, ofrecen una utilización eficiente de estos combustibles.

Las células de combustible pueden clasificarse por el electrolito, por la temperatura de operación, por el oxidante y por el combustible. Podría suponerse una combinación ilimitada de estos factores, pero son relativamente pocas las combinaciones prácticas. Esta limitación se debe a las restricciones que el electrolito impone en los otros factores, y viceversa. Así, la elección de ácido fosfórico como electrolito establece un intervalo de temperaturas de operación desde 150 a 200 grados Celsius. Por debajo de 150 grados Celsius, el ácido fosfórico tiene una conductividad baja; por encima de 200 grados Celsius, los materiales de los electrodos son inestables. Una célula de combustible con un electrolito de carbonato fundido requiere una temperatura de operación de 600 a 700 grados Celsius y una corriente de combustible que contenga óxidos de carbono, así como también hidrógeno.

Las células que funcionan con hidrógeno potásico como electrolito operan mejor entre 50 y 150 grados Celsius. Este electrolito no tolera óxidos de carbono ni en el combustible ni en el oxidante. La reacción del hidrógeno con los óxidos de carbono conduce a la formación de carbonato potásico, que rebaja drásticamente el rendimiento de la célula. Las células de combustible con este electrolito (y oxígeno e hidrógeno puros como reactantes) se han utilizado para ingenios espaciales, pero su uso no espa-



FUENTES DE COMBUSTIBLE para la generación de energía eléctrica, comparadas para los diferentes tipos de generadores en 1976 *(gris obscuro)*, 1985 *(color)* y 2000 *(gris claro)*. Una planta generadora de electricidad de carga básica *(barras de la izquierda)* opera continuamente y funciona mejor con energía nuclear o carbón. Una planta siguiendo la carga *(barras del centro)* funciona durante el día, excepto los fines de semana, siguiendo las oscilaciones de la carga, y una planta para atender los máximos de carga *(barras de la derecha)* se pone en marcha sólo cuando la demanda de energía eléctrica es alta. El combustible para las plantas siguiendo la carga o de atención de picos será combustible líquido o gaseoso de origen fósil. Las plantas de células que se alimentan con combustibles líquidos se utilizarían como plantas siguiendo la carga o de atención a los máximos de carga.

cial está limitado por el coste de obtención de reactantes libres de óxidos de carbono.

Las células de combustible con ácido sulfúrico, ácido sulfónico o polímeros sólidos como electrolito están limitadas por la presión de vapor del agua en el electrolito. Si el aire es la fuente del oxígeno, estas células deben operar a temperaturas por debajo de los 100 grados Celsius; de otro modo, se iría demasiada agua con el nitrógeno (el otro constituyente importante del aire) que se purga, y el electrolito se haría inestable. Sin embargo, a temperaturas por debajo de los 100 grados Celsius no puede conseguirse la elevada eficacia deseable para una central energética de células de combustible que suministre electricidad. Ello explica que no se preste demasiada atención a estos electrolitos.

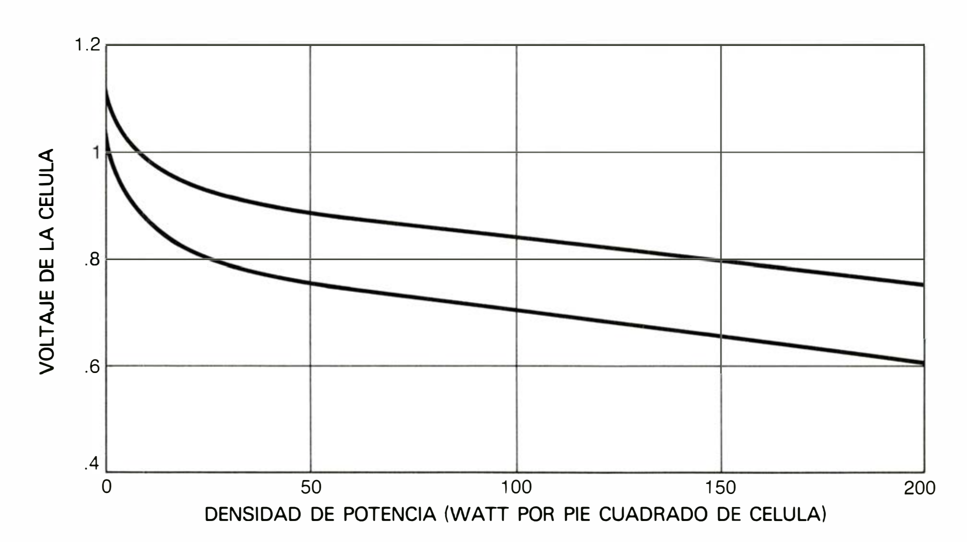
Los electrolitos que están estudiándose en la actualidad son el ácido fosfórico y las sales carbonatadas fundidas. El año pasado se gastaron aproximadamente unos 48 millones de dólares por las oficinas gubernamentales de los Estados Unidos y organizaciones como el Electric Power Research Institute en el estudio de células de combustible no espaciales. Unos 39 millones de dólares de dicha cantidad se han invertido en la investigación de células de ácido fosfórico y unos 7 millones de dólares en células de carbonatos fundidos.

A lo largo de varios años se han venido realizando también investigaciones sobre electrolitos sólidos que conducen iones de oxígeno. Esta célula de combustible de óxidos sólidos opera a temperaturas de unos 1000 grados Celsius. La tecnología tiene todavía un considerable trabajo a desarrollar antes de que puedan emplearse en células de combustible con rendimiento útil.

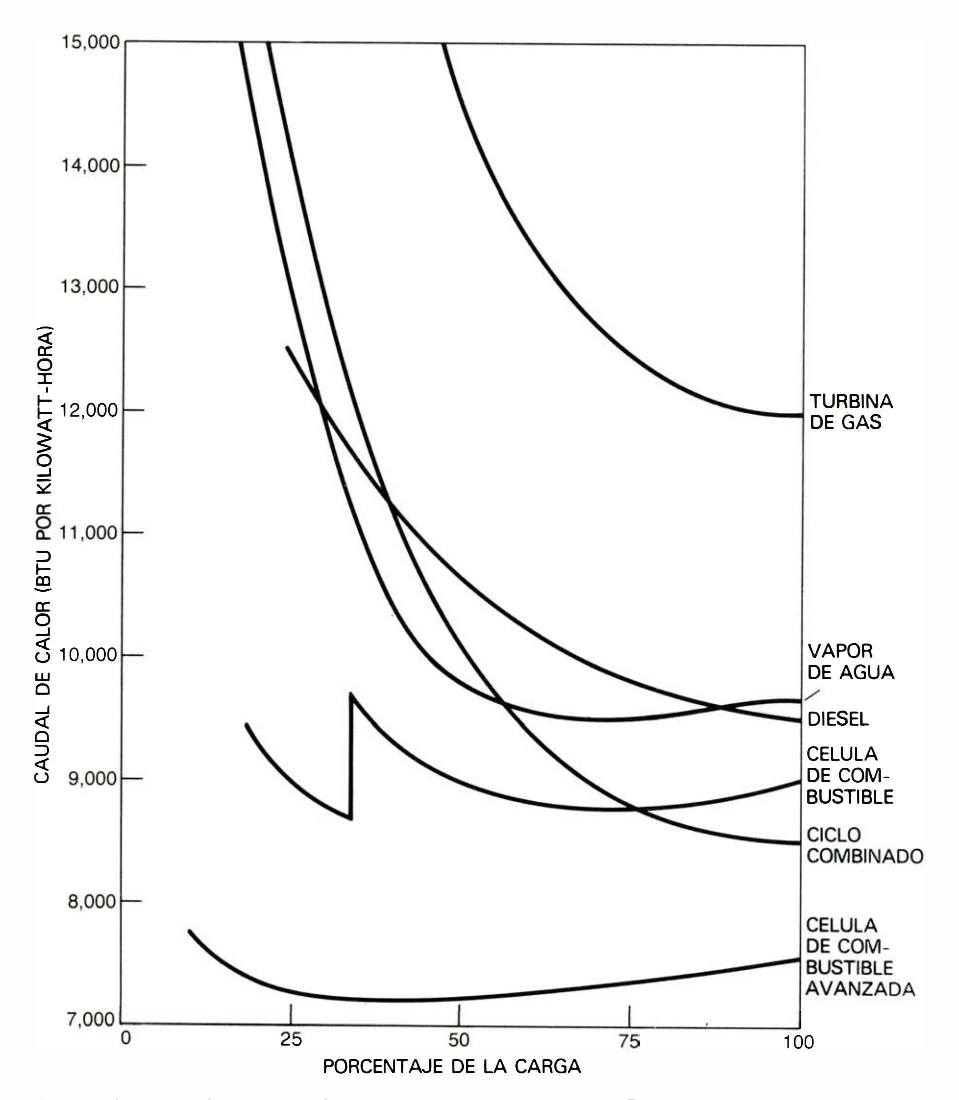
El interés de la célula de combustible como central energética de suministro procede de su eficacia, su aceptabilidad respecto al medio ambiente y su configuración modular. En lo que respecta a la eficacia el punto clave es que, al no tratarse de una máquina térmica, la célula de combustible no está limitada por el ciclo de Carnot, que describe los límites de la eficacia de las máquinas térmicas. Por tanto, la célula de combustible ofrece la posibilidad de lograr eficacias de conversión mayores que las que pueden conseguir los generadores térmicos. Puesto que la eficacia de las células de combustible está determinada principalmente por el voltaje de una célula única, la eficacia de una central de células de combustible es (hasta cierto grado) independiente del tamaño de la planta.

Además, en contraste con los generadores convencionales, la eficacia de la célula de combustible es aproximadamente constante en el intervalo de 25 a 100 por ciento de su potencia energética. Por tanto, las células de combustible que proporcionan energía siguiendo la carga ofrecen la posibilidad de conservar combustible. Esta capacidad puede visualizarse mejor considerando el caudal de calor en vez de la eficacia. El caudal de calor es una expresión de la cantidad (expresada en Unidades Térmicas Británicas: Btu) de combustible necesario para generar un kilovatt-hora de electricidad. Para una potencia energética determinada, una central energética que utiliza petróleo y opera en un ciclo combinado (en el que el calor residual de un generador primario, una turbina de gas por ejemplo, se emplea para hacer funcionar un generador secundario, una turbina de vapor) constituye el generador energético más eficaz que podamos ahora disponer; requiere 8500 Btu de combustible por kilovatt-hora, que hemos de comparar con la necesidad proyectada de 9000 Btu para una central energética basada en células de combustible. Sin embargo, para un 40 por ciento de la capacidad, que es una cifra realista en operaciones siguiendo la carga, la central de ciclo combinado requerirá 11.250 Btu por kilovatt-hora, que debemos comparar con 9250 para una central de células de combustible. Por consiguiente, la célula de combustible ofrece el mayor potencial para la conservación del combustible cuando se utiliza siguiendo la carga, mientras que son aconsejables los otros generadores para operar en condiciones cercanas a su eficacia máxima.

Puesto que la célula de combustible produce energía mediante un proceso electroquímico y no por combustión, las únicas emisiones distintas de aire y agua serán las procedentes del tra-



CARACTERISTICAS DE RENDIMIENTO de una célula de combustible. Tienen una gran influencia sobre el diseño y coste de una central energética. Se han representado la densidad de energía y el voltaje para una célula de combustible con un electrolito de ácido fosfórico *(curva inferior)* y como se proyectan para una célula de combustible con un electrolito de carbonato fundido *(curva superior)*. El coste de una "sección energética" que genera energía a partir de células de combustible disminuye a medida que la densidad de potencia aumenta; pero también lo hace el voltaje. El diseño de plantas consiste con frecuencia en ajustar el voltaje al nivel deseado de eficacia y aceptar la densidad de potencia resultante. Los esfuerzos se centran en obtener un voltaje superior para una densidad de potencia determinada.



POTENCIAL DE CONSERVACION de la célula de combustible. Puede verse comparando su caudal de calor (número de unidades térmicas británicas, Btu, de combustible para producir un kilowatt-hora de energía eléctrica) con el de otros tipos de plantas en un amplio intervalo de cargas. Las instaladas para funcionar siguiendo la carga permitirían a otros generadores operar a su mayor eficacia.

tamiento del combustible. Las centrales energéticas de células de combustible que operan con combustibles procedentes del petróleo o carbón producirán emisiones de dióxido de azufre y óxido nitroso de menos de 0,045 y 90 gramos por un millón de Btu, respectivamente. Estas emisiones y otras son 10 veces, por lo menos, inferiores que los límites estándar que rigen para todo el territorio de los Estados Unidos.

Las emisiones de azufre son bajas porque las centrales energéticas de células de combustible no toleran bien los compuestos de azufre; y así, estos compuestos se separan mediante un subsistema especial de la planta de tratamiento de combustible. Las otras emisiones son bajas debido a la limpieza inherente de los procesos electroquímicos. Una central de células de combustible opera silenciosamente y no requiere aportación de agua a la temperatura ambiente de 35 grados Celsius o más bajas. Merced a estas características positivas en relación con el medio ambiente, puede instalarse una central de células de combustible en la misma zona urbana a la que alimenta. Distribuyendo las centrales de células de combustible, la compañía de suministro puede evitar inversiones en nuevas líneas para la transmisión y distribución de la energía y puede reducir las pérdidas inherentes a las líneas de transmisión.

La célula de combustible es de configuración esencialmente modular. Ello permite que los componentes de una planta puedan fabricarse en una factoría mediante técnicas de producción en serie en vez de tener que construirla in situ. Por tanto, los costes de construcción serán inferiores. Además, la disponibilidad de pequeños generadores modulares permite a una compañía añadir bloques de capacidad relativamente pequeños, igualando el crecimiento de la demanda de una forma económica. La existencia de pequeñas plantas modulares seguras aumentaría también la seguridad inherente de un sistema de suministro y facilitaría una reducción en la capacidad generativa de reserva que la compañía mantiene.

Como consecuencia de estas características, que no se encuentran en los generadores energéticos convencionales, la célula de combustible podría abastecer una red de suministro de diversas maneras. Podrían montarse centrales energéticas, relativamente pequeñas, de células de combustible (con un intervalo de capacidad de 25-200 kilowatt) en edificios comerciales y residenciales. Una planta de este tipo podría utilizar gas natural

como combustible. La planta proporcionaría energía eléctrica y térmica (esta última del calor residual de las células de combustible), consumiendo la misma cantidad de combustible requerido ordinariamente para la demanda térmica sólo. Se han proyectado eficacias próximas al 100 por ciento para centrales energéticas de células de combustible de este tipo. Las plantas tendrían la ventaja adicional de alargar el suministro de gas natural.

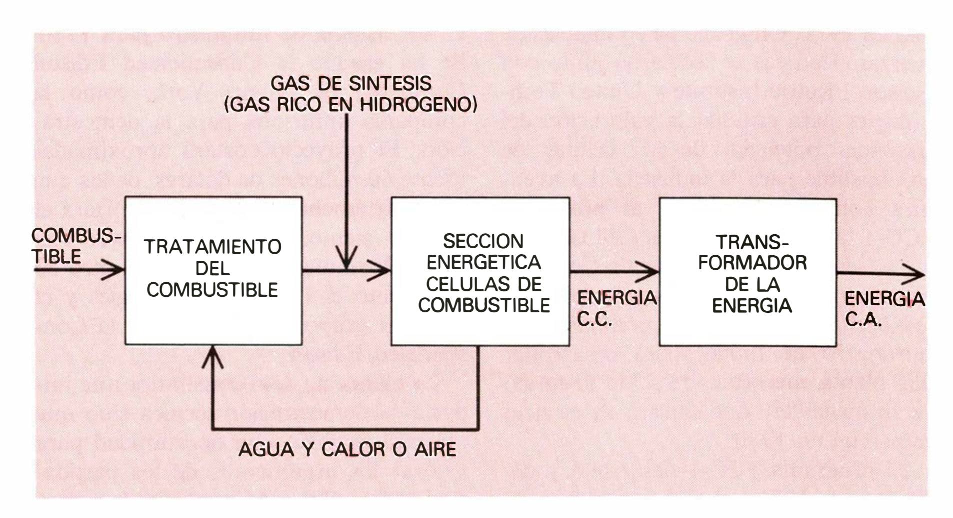
Se podrían distribuir plantas mayores, con un intervalo de capacidad desde 5 a 25 megawatt, a través de un sistema de suministro eléctrico para llevar a cabo eficazmente la operación siguiendo la carga. Si se equiparan para recuperar el calor residual (materia que está en estudio) representarían un uso final altamente eficaz de los recursos de combustibles fósiles. Son alcanzables eficacias superiores al 80 por ciento (basadas en el consumo de combustible).

En un futuro más lejano, las células de combustible podrían integrarse con gasificadores de carbón para constituir grandes centrales energéticas que utilizaran carbón directamente. La capacidad de estas plantas variaría desde 150 a 600 megawatt. Una planta de este tipo está proyectada para una eficacia superior al 45 por ciento, basada en el valor calorífico del carbón consumido.

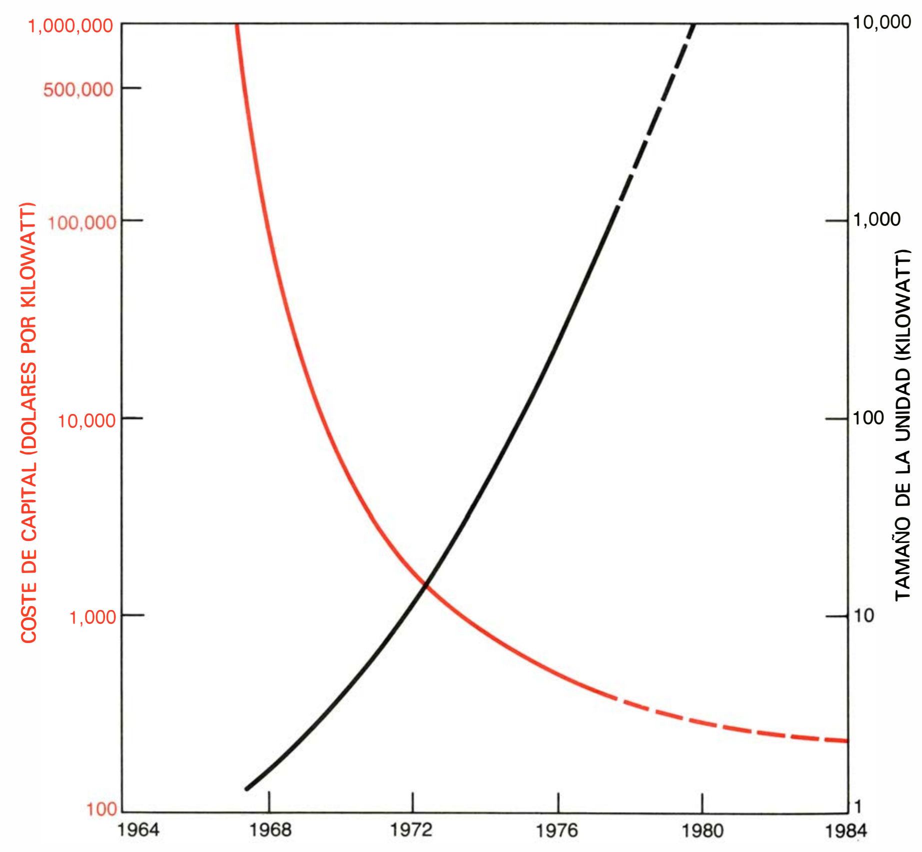
Hay tres proyectos de primera línea que se aprestan a acelerar el desarrollo de la célula de combustión y una tecnología comercial concomitante. Uno es una extensión de un proyecto denominado TARGET (acrónimo de "Team to Advance Research for Gas Energy Transformation", Equipo para Fomentar la Investigación en la Transformación de Energía de los Gases). El proyecto fue patrocinado en un principio por la United Technologies Corporation juntamente con un consorcio de compañías de gas y de obtención de electricidad a partir de gas. Desde su iniciación, en 1967, el grupo ha venido financiando el desarrollo de células de combustible de ácido fosfórico para aplicaciones residenciales y comerciales in situ.

El esquema prototípico de centrales energéticas TARGET utiliza gas (natural o sintético) como combustible, convirtiéndolo en electricidad y en energía térmica para calefacción y refrigeración. La eficacia es bastante elevada, ya que virtualmente todo el calor residual se dirige hacia la parte de calefacción y refrigeración de la operación. Hasta ahora la industria ha gastado unos 100 millones de dólares en el proyecto. En 1972 y 1973 se ensayaron para las condiciones in situ más de 60 plantas energéticas de 12,5 kilowatt de capacidad: los ensayos condujeron al desarrollo y demostración de una planta de 40 kilowatt en 1975. Ahora, en una ampliación del programa con ayuda del Departamento de Energía y del Gas Research, el objetivo es experimentar con 50 plantas de 40 kilowatt de capacidad, entre 1979 y 1981, y la obtención de la viabilidad comercial de la tecnología en 1982.

El segundo proyecto vino de la mano de la industria de suministro de electrici-



CENTRAL DE ENERGIA basada en células de combustible, representada esquemáticamente. Un combustible, que podría ser carbón o gas o un destilado del petróleo, se convierte, a través de un tratamiento del combustible, en un gas rico en hidrógeno, que es el que entra en las células de combustible. La energía generada por una sección energética de células de combustible es corriente continua, que debe convertirse mediante un transformador de energía en corriente alterna para su distribución.



COSTE DE CAPITAL Y TAMAÑO DE LA PLANTA comparados de las centrales energéticas de células de combustible existentes y proyectadas. La progresión en el tamaño unidad *(negro)* y el coste de capital *(color)* va desde una unidad de 1,5 kilowatt, que fue construida para el ejército estadounidense a finales de los años 60 a un coste que excedía los 100.000 dólares por kilowatt, hasta la unidad de 4,5 megawatt, que se ensayará en Nueva York a principios de los años 80. Las líneas sólidas reflejan las cifras reales, las líneas a trazos las cifras proyectadas. Los costes son proyecciones para la producción.

dad en 1971. Un grupo de compañías de fuerzas eléctricas se asociaron junto con Edison Electric Institute y United Technologies para estudiar la valoración del beneficio potencial de las células de combustible para la industria. La aventura condujo, en 1972, al programa FCG-1 (acrónimo de "Fuel-Cell Generator", Generador de Células de Combustible), empresa subvencionada por United Technologies y nueve compañías de suministro de fluido, para desarrollar una planta energética de ácido fosfórico de 26 megawatt, que entrará en servicio comercial en 1980.

El programa FCG-1 desarrolló y demostró (en 1976 y 1977) una planta piloto de un megawatt. La demostración reveló que una planta energética, que consuma nafta, proporcionará grandes cantidades de energía a un sistema de suministro, reuniendo al mismo tiempo satisfactoriamente los requisitos operacionales de la industria de caudal de calor, emisiones y capacidad de seguir la carga.

En 1976, con el propósito de acelerar la aplicación comercial de las células de combustible, el Electric Power Research Institute y la Energy Research and Development Administration (ahora Departamento de Energía) entraron a formar parte del programa FCG-1. De ello surgió el proyecto de diseñar, construir y someter a ensayo un módulo de 4,5 megawatt de la planta energética FCG-1 en un sistema de suministro para 1980. Se ha elegido la Consolidated Edison Company, de Nueva York, como la compañía anfitriona para la demostración. El proyecto costará aproximadamente 60 millones de dólares, de los que el Departamento de Energía aportará el 55 por ciento, el Electric Power Research Institute el 20 por ciento, otro 20 por ciento el United Technologies y el resto un grupo encabezado por la Consolidated Edison.

La planta no sólo constituirá una importante demostración técnica sino que ofrecerá también una oportunidad para valorar los argumentos de los partidarios de las células de combustible acerca de las ventajas de esta tecnología en lo tocante a versatilidad y facilidad de asentamiento. La planta estará en la parte baja de la ciudad de Nueva York, donde no se podría colocar un generador convencional de combustible fósil debido a sus emisiones. Localizadas en áreas urbanas, las centrales energéticas, limpias y dispersas, ofrecen una de las dos opciones para aumentar la capacidad de generación eléctrica al objeto de satisfacer la creciente demanda. (La otra opción es construir plantas convencionales en lugares remotos y transportar la electricidad hasta la ciudad. El coste de obtención de derechos de paso y de construcción de nuevas líneas de transmisión se ha convertido en algo casi prohibitivo.)

El tercer gran proyecto surgió de los esfuerzos emprendidos por industrias de suministro de gas y electricidad con objeto de mejorar la tecnología de las células de combustible. El programa, asumido ahora también por el Departamento de Energía y el Electric Power Research Institute, pretende construir una planta energética de células de combustible de carbonato fundido que estará alimentada por combustibles derivados del carbón y se hallará disponible para usos comerciales dentro de unos diez años. Las organizaciones implicadas en el programa son el Argonne National Laboratory, la Energy Research Corporation, la General Electric Company, el Institute of Gas Technology y el Oak Ridge National Laboratory.

La tecnología de la célula de combustible de ácido fosfórico está muy desarrollada. Cada célula consta de una matriz porosa delgada que contiene ácido fosfórico concentrado. Esta matriz de electrolito está situada entre electrodos de carbón poroso catalizados con pequeñas cantidades de platino, típicamente de 2,6 a 28,9 gramos por metro cuadrado de área de célula. Las células trabajan a temperaturas de 150 a 190 grados Celsius y producen de 1000 a 2000 watt de potencia por metro cuadrado a 0,64 volt de corriente continua (suponiendo que la eficacia de la planta sea del 38 por ciento).

PROYECTO	FUNCION	FUNCIONAMIENTO	TAMAÑO	COMBUSTIBLE	COSTE	EFICACIA	DISPONIBILIDAD	ELECTROLITO
TARGET	PLANTAS ENERGETICAS IN SITU	SIGUIENDO LA CARGA (CONTINUO)	25 A 200 KILOWATT	GAS NATURAL	500	35 A 40 O 100	A PRINCIPIOS DE LA DECADA 1980	ACIDO FOSFORICO
FCG-1	PLANTAS ENERGETICAS DISPERSAS	SIGUIENDO LA CARGA (INTERMEDIO)	10 A 25 MEGAWATT	NAFTA	350	37 A 39 O 75 A 80	A PRINCIPIOS DE LA DECADA 1980	ACIDO FOSFORICO
SUMINISTRO	PLANTAS ENERGETICAS DISPERSAS	SIGUIENDO LA CARGA (INTERMEDIO)	10 A 25 MEGAWATT	PETROLEO O LIQUIDOS DERIVADOS DEL CARBON	350	45 A 47 O 75 A 80	A FINALES DE LA DECADA 1980	CARBONATO FUNDIDO O ACIDO FOSFORICO
SUMINISTRO	PLANTAS ENERGETICAS CENTRALES	CARGA BASICA	150 A 600 MEGAWATT	CARBON	800	45 A 47	AÑOS 1990	CARBONATO FUNDIDO O ACIDO FOSFORICO

PRINCIPALES PROYECTOS de células de combustible. El programa TARGET (acrónimo de Team to Advance Research for Gas Energy Transformation, Equipo para Fomentar la Investigación para la Transformación de la Energía de los Gases) fue acometido por la industria de suministro de gas y el programa FCG (acrónimo de Fuel-Cell Generator, Generador de Células de Combustible) por la industria de suministro eléctrico. Los proyectos denominados "Suministro" representan planes a largo plazo de la industria de suministro eléctrico. Las cifras indicadas en "Coste" son centenares de dólares (valor 1980) por kilowatt de capacidad instalada. La eficacia se da en porcentaje; donde se presentan cifras alternativas, las de la línea superior son plantas que operan sin recuperación del calor residual y las de la línea inferior son para plantas con recuperación del calor.

La actual célula de ácido fosfórico resulta casi adecuada para aplicaciones de suministro. Se ha demostrado su eficacia para densidades de energía que están de acuerdo con costes razonables del conjunto, siempre que los conjuntos se fabriquen mediante técnicas de producción en serie. El coste de capital no está controlado por el coste inherente a los materiales; ahora, el factor determinante es el nivel de producción. El Electric Power Research Institute estima que podrían producirse en serie plantas energéticas de ácido fosfórico completas por 350 dólares por kilowatt (dólares de 1980) si el ritmo de producción fuera de 500 megawatt por año. Esto podría concordar con el coste de otros generadores diseñados para operaciones siguiendo la carga.

Aunque pequeños conjuntos de células de ácido fosfórico han operado durante muchos miles de horas, las 40.000 horas (cuatro años y medio) de seguridad deseadas mediante un gran conjunto están todavía por demostrarse. La planta de demostración de 4,5 megawatt, que ha de construirse en la ciudad de Nueva York, está planeada para funcionar durante 6700 horas. Las mejoras de la tecnología incorporada en la demostración de Nueva York prometen conjuntos de células que conseguirán las necesarias 40.000 horas. La vida y seguridad de las plantas de ácido fosfórico se establecerán después de haber adquirido más pruebas con plantas grandes. Pasada esta primera generación de células de combustible de ácido fosfórico, las perspectivas son excelentes para avances evolutivos que prometen eficacias mayores y costes inferiores.

La tecnología del carbonato fundido lleva, como mínimo, cinco años de retraso con respecto a la del ácido fosfórico. El electrolito es una delgada capa compuesta de carbonato fundido y un relleno inerte; está colocado entre dos electrodos, hechos de níquel poroso. A la temperatura de operación de 650 grados Celsius la velocidad de reacción en los electrodos es tan elevada que no se necesita ningún catalizador especial. En las condiciones de operación se forma óxido de níquel en el electrodo del oxígeno (cátodo) y se convierte en el material catódico activo; el níquel constituye el material anódico activo.

Las células de carbonato fundido han manifestado una densidad de potencia de 1000 a 1250 watt por metro cuadrado de área celular a 0,785 volt (eficacia de la planta energética del 45 por ciento). Este voltaje superior y la tolerancia a los óxidos de carbono constituyen las ventajas de la tecnología del carbonato fundido. Una pequeña célula ha funcionado durante 14.000 horas y un conjunto formado por 19 células de unos 1000 centímetros cuadrados cada una ha estado funcionando durante 1400 horas.

No obstante estos resultados prometedores, quedan problemas importantes por resolver. Se refieren, sobre todo, a los cambios estructurales que tienen lugar en el electrolito y a la pérdida gradual del mismo en períodos largos de operación. Tales problemas limitan la resistencia y seguridad de las células, pero existen razones para esperar que se puedan solucionar. La mejor estima del coste de una planta energética de carbonato fundido conseguida en producción en serie es de 350 dólares por kilowatt-hora (dólares de 1980), el mismo que una planta de ácido-fosfórico.

Los otros subsistemas –tratamiento del combustible y acondicionador de la energía– están virtualmente preparados ya para la primera generación de plantas energéticas de células de combustible. En otras palabras, se dispone de sistemas de tratamiento del combustible capaces de reformar gas natural y nafta de modo que puedan funcionar con una planta de células de combustible. Se necesitarán esfuerzos ulteriores para ampliar la capacidad de procesado de combustibles de nafta a líquidos más pesados procedentes del petróleo y líquidos obtenidos a partir del carbón. Los acondicionadores de energía ahora disponibles parecen ser satisfactorios tanto para la primera generación de plantas de células de combustible como para las más avanzadas que vendrán posteriormente.

El objetivo a largo plazo de integración de un sistema de células de combustible y gas obtenido a partir de carbón requerirá el desarrollo de gasificadores económicamente interesantes. Existen en estudio muchos proyectos importantes de gasificación. En efecto, el esfuerzo centrado en el estudio de la gasificación del carbón en los Estados Unidos es considerablemente mayor que el esfuerzo dedicado a las células de combustible.

Los beneficios que la célula de combustible ofrece a la industria de suministro y a la sociedad se obtendrán sólo después de que la célula de combustible alcance la escala comercial. La comercialización es actualmente el principal obstáculo al que se enfrentan las células de combustible. Otro problema es el enorme grado de incertidumbre de la política energética en los Estados Unidos. Una cuestión más importante es el precio de las células de combustible, que no alcanzará un nivel que sea aceptable para los usuarios potenciales hasta que se consiga la fabricación de un tamaño adecuado y se hayan producido un gran número de plantas energéticas.

El valor de las células de combustible de multimegawatt para el suministro eléctrico variará desde 330 dólares hasta quizá 600 dólares por kilowatt, dependiendo de otras limitaciones de suministro y de la evaluación de los beneficios indirectos de las células de combustible tales como la ausencia de inversión en nuevas líneas de transmisión y distribución y la reducción de los márgenes generativos de reserva. Es probable que lleguen a producirse células de combustible entre 500 y 1500 megawatt antes que el precio alcance el nivel objetivo de 350 dólares por kilowatt. Para un mercado espontáneo de escala comercial a desarrollar es evidente que debe darse alguna combinación de los tres factores siguientes: (1) el fabricante debe asumir el riesgo de suscribir la variación desde el nivel de 1500 dólares por kilowatt, que es lo que la primera unidad de producción puede costar, hasta el nivel objetivo de 350 dólares; (2) los usuarios deben comprar las primeras unidades a su coste real (1500 dólares por kilowatt, decreciendo las subsiguientes unidades en coste hacia el nivel último de 350 dólares); y (3) el Gobierno debe asistir al sector privado reduciendo los riesgos o ayudando a sufragar el coste de las unidades iniciales.

Las incertidumbres acerca de la política nacional sobre energía, del suministro de combustible, de la demanda futura de fluido eléctrico y de la estructura de los suministros eléctricos, unidas a las incertidumbres inherentes a cualquier tecnología nueva e inmadura impedirán, probablemente, que los fabricantes y las compañías eléctricas emprendan el riesgo que comportaría llevar la célula de combustible a nivel comercial. Puesto que los beneficios de la célula de combustible se extienden más allá del sector privado a la sociedad en general, conviene que el Gobierno asista al sector privado en su desarrollo. La célula de combustible es quizá la primera nueva tecnología de la energía que ha emergido en las últimas dos décadas y puede, por tanto, soportar la carga de explorar mecanismos de comercialización. No es del todo imposible que la célula de combustible sea abandonada debido a la incapacidad de los sectores público y privado de tratar esta novedad compleja de manera oportuna.

Ciencia y sociedad

Pistas rápidas

En octubre de 1977, la Universidad de Harvard inauguró un nuevo polideportivo cubierto, que contiene una pista para carreras, de seis calles y 203 metros de longitud. La infraestructura de la pista está hecha principalmente de madera y de un revestimiento superficial sintético. La mayoría de quienes la han utilizado dicen que correr por ella constituye una experiencia agradable e insólita. Se la siente singularmente elástica. Un análisis de las marcas de los corredores de Harvard y de sus competidores en 1977-1978 muestra que la pista mejora notablemente el rendimiento de un corredor. Thomas A. McMahon y Peter R. Greene han diseñado un modelo de la mecánica de la carrera atendiendo a las interacciones músculo-pista.

En sus experimentos, unos levantadores de pesas que llevaban sobre sus hombros cargas hasta del doble de su peso corporal ejecutaban pequeños movimientos de rebote sobre una larga tabla, manteniendo las rodillas dobladas hasta el grado en que lo harían al correr. Conociendo la masa total del sujeto y el peso (y, por tanto, la fuerza ejercida sobre los músculos), la masa efectiva de la tabla y la frecuencia natural de vibración, emplearon un procedimiento matemático sencillo para calcular la rigidez de los sujetos. Para ángulos de flexión en la rodilla del orden apropiado para correr, la rigidez controlada, o refleja, es casi independiente de la carga y, por tanto, de la fuerza muscular.

La observación acerca de que los reflejos mantienen la rigidez muscular a un nivel casi constante resultó tener gran valor práctico. Los resultados de los experimentos implican que el cambio que sobreviene en la rigidez del músculo con la fuerza muscular, observado con experimentos con músculos aislados, puede despreciarse. De hecho, durante aproximadamente el primer cuarto del período de contacto con el suelo de un paso de carrera, los reflejos no han tenido tiempo de actuar, de modo que sólo las propiedades intrínsecas de los músculos determinan la rigidez y el amortiguamiento de la pierna. Pero en la segunda mitad del período de contacto, la rigidez y el amortiguamiento están casi bajo control reflejo y quedan, por tanto, razonablemente bien representados por un resorte lineal amortiguado.

Los músculos y reflejos de la pierna pueden verse como un sistema compuesto de un elemento de piñón y cremallera que funciona en serie con un resorte amortiguado. El elemento de piñón y cremallera destaca la función de las órdenes de movimiento que desciende de los centros motores superiores; el resorte representa la supuesta rigidez constante de músculos y reflejos, y el mecanismo amortiguador representa la propiedad intrínseca de fuerza-velocidad de los músculos y la sensibilidad de los reflejos a la velocidad de cambio. En otras palabras, la fijación de las posiciones de la pierna viene modelada por el movimiento del sistema piñón-cremallera, y el efecto de perturbaciones ocasionadas por fuerzas externas, como el impacto de la pista sobre la pierna del corredor, está modelado por la flexión del resorte amortiguado. Las simplificaciones introducidas por el modelo hacen posible calcular la velocidad de un corredor en una pista de elasticidad arbitraria contando sólo con pocos elementos de información: la masa del corredor, la longitud de su pierna, su rigidez de resorte, la longitud de su paso y su máxima velocidad sobre una superficie dura.

¿Qué valor de elasticidad de pista —si lo hay— optimizará la velocidad de carrera? Para nuestros fines, la manera más útil de medir la velocidad de carrera consiste en hacerlo por la relación entre longitud de paso (la distancia que recorre el cuerpo hacia adelante mientras un pie está aún tocando el suelo) y el tiempo de contacto en el suelo (el tiempo que permanece el pie en el suelo). En una pista ajustada, debería minimizarse el tiempo de contacto y maximizarse la longitud de paso. En el modelo, lo que determina primariamente el movimiento del cuerpo durante la fase de contacto en el suelo, y, por tanto, el tiempo de contacto con el suelo, son la masa del corredor y el movimiento del resorte amortiguado. El movimiento del elemento cremallera-piñón determina las posiciones de la pierna y la cadera al inicio y al término de la fase de contacto con el suelo, y, por tanto, la longitud del paso.

Examinemos primero la relación entre rigidez de la pista y tiempo de contacto con el suelo. Para deducir una expresión del tiempo de contacto con el suelo como función de la rigidez de la pista es preciso asignar a un corredor una rigidez de resorte y determinar también hasta qué grado los efectos conjugados con la relación fuerza-velocidad y de la realimentación por velocidad, realimentación procedente de los reflejos, actúan amortiguando su resorte. Hemos observado que una vez que los músculos de un corredor han guiado su pie hasta que hace contacto con el suelo, el papel más importante de los músculos es contraerse con la máxima fuerza y brevedad a fin de poder invertir la velocidad descendente del cuerpo. Por este motivo optaron por hacer la suposición, algo arbitraria, de que el sistema de cremallera y piñón no podía ser perturbado por fuerzas exteriores mientras el pie estuviera en el suelo; así, en una superficie dura, el tiempo de contacto con el suelo depende íntegramente de la masa del corredor y de las propiedades del resorte amortiguado. Cuando corre sobre una superficie dura, a partir de la medición de su tiempo de contacto con el suelo pueden deducirse los valores numéricos que denotan la rigidez y el amortiguamiento del resorte de un corredor.

Determinadas las constantes de resorte y rigidez, puede calcularse cómo cambia el tiempo de contacto con el suelo en pistas de diferente rigidez. Este resultado fue el más importante para predecir el efecto de la rigidez de la pista sobre la velocidad de carrera. Puede suponerse que el corredor está ajustado a la pista de modo tal que rebota desde ella en medio período del modo más bajo de vibración natural. Tal como lo sugiere la intuición, los cálculos muestran que el tiempo de contacto en superficies blandas es prolongado. Si en el resorte del músculo humano no hubiese elemento amortiguador, el tiempo de contacto con el suelo decrecería continuamente desde ese valor conforme disminuye la elasticidad, alcanzando un mínimo a la elasticidad de superficies infinitamente duras.

En tal caso, una superficie para correr como la del hormigón sería la más rápida. Pero los resultados muestran que el tiempo de contacto con el suelo no decrece a todo lo ancho de la gama de elasticidades. Por ser el corredor un sistema amortiguado, su tiempo de contacto alcanza un mínimo a una elasticidad intermedia; a una rigidez de pista de dos a cuatro veces mayor que la del corredor, el tiempo de contacto es más breve que en superficies más duras. Por consiguiente, en pistas con una rigidez precisamente correcta serían posibles velocidades mejoradas.

Pero la velocidad de carrera no está enteramente determinada por el tiempo de contacto con el suelo; la longitud del paso constituye también un factor importante. Recuérdese que longitud de

paso es la distancia recorrida mientras el pie está en el suelo, en contraposición a la distancia entre huellas plantares consecutivas, que es la longitud de zancada. Cuando se corría sobre las pistas experimentales, más elásticas, como la construida con bloques de caucho alveolar, la longitud de los pasos era mayor. La explicación de este fenómeno tiene relación con la posición de la pierna de contacto al inicio y al término de la fase de contacto. En esos puntos la pierna está totalmente extendida, y puede suponerse que no hay cargas que actúan sobre el resorte; es decir, la geometría de la pierna, conforme golpea el suelo, depende únicamente del elemento cremallera-piñón.

En las investigaciones del movimiento del sistema piñón-cremallera se adoptó otro supuesto basado en un aspecto que revelaron las películas tomadas de las carreras. Por dura o blanda que fuese la superficie de la pista, la trayectoria de la cabeza de todos los corredores era casi horizontal, como si sus reguladores de postura estuviesen intentando dar al mecanismo detector de aceleración, existente en el oído medio, un viaje sin sobresaltos. En cualquier caso es indudable que, a la frecuencia en cuestión, con el pequeño valor de aceleración debido a la gravedad y el breve tiempo entre una y otra pisada, el movimiento vertical del centro de masa del cuerpo no puede ser grande. Por tanto, se optó por suponer que la cadera sigue un recorrido horizontal cuando el corredor corre la pista, sea ésta dura o blanda.

La longitud del paso sobre una pista de elasticidad arbitraria puede calcularse del modo siguiente. En cada paso de carrera, una de las piernas del corredor entra en contacto con el suelo, mientras la otra oscila sobre la superficie de la pista. Conforme la pierna oscilante se desplaza hacia adelante, permanece por encima de la superficie no flexionada de la pista, sea ésta dura o blanda. Sobre una pista dura, la pierna de contacto permanece también por encima de la superficie. Pero en una pista blanda, la pista de contacto se hunde bajo la superficie. Por consiguiente, es necesario visualizar al corredor moviéndose a través de una superficie hipotética por una distancia δ por debajo de la superficie no flexionada, en donde δ es la flexión media de pista a lo largo de un ciclo completo de zancada. (El ciclo de zancada incluye no sólo la fase de contacto con el suelo sino también la fase aérea: si el corredor estuviese inmóvil, de pie, sobre la pista que cede, estaría de pie en un pozo de profundidad δ.)

Al comienzo (y al final) de la fase de contacto, la pierna de contacto del corredor está totalmente extendida, y a la distancia entre cadera y talón la llamaremos l. A medio camino, cuando se ha recorrido la distancia de medio paso, la pierna de contacto se flecta, acortándose en, digamos, una longitud δ_0, sobre una pista dura, de modo que la distancia entre cadera y superficie de pista dura es $l-\delta_0$. Suponiendo que la cadera sigue un recorrido horizontal, esta situación queda descrita por un triángulo rectángulo de hipotenusa l y lados $l-\delta_0$ y $L_0/2$, en donde L_0 es la longitud del paso en una superficie dura. Pero en una pista elástica, el triángulo correspondiente tiene lados l, $L/2$ y $l-\delta_0-\delta$, en donde L es la longitud del paso y $l-\delta_0-\delta$ es la distancia desde la cadera hasta la superficie no flexionada de la pista. La flecha δ está determinada por dos cantidades mensurables: el peso corporal del corredor y la rigidez de la pista. La longitud de la pierna, l, y la longitud del paso sobre una pista dura, L_0, pueden medirse directamente. Por tanto, aplicando el teorema de Pitágoras a ambos triángulos, es asunto sencillo calcular cuál será la longitud del paso de un corredor sobre una pista de rigidez arbitraria.

La longitud del paso se incrementa al máximo (efecto deseable) en la gama de elasticidades de pista en las que se incrementa el tiempo de contacto con el suelo (efecto indeseable). Pero resulta que en la gama de elasticidades que son óptimas respecto al tiempo de contacto con el suelo, la longitud de paso aún se incrementa ligeramente por sobre su valor en superficie dura. Por consiguiente, en esta gama de elasticidades a una rigidez de poco más de doble de la rigidez del corredor, los dos factores que determinan la velocidad de carrera se aúnan de manera óptima. Los resultados del modelo mostraban que el efecto conjugado del menor tiempo de contacto y de la mayor longitud del paso llevarían, en pistas cuya elasticidad estuviese dentro de esa gama, a velocidades incrementadas del 2 al 3 por ciento.

Política sanitaria

En febrero de 1976, el Centro para Control de Enfermedades del Gobierno Federal de los Estados Unidos identificó el agente responsable de un pequeño brote de una enfermedad del aparato respiratorio entre los reclutas de Fort Dix, Nueva Jersey, como un subtipo particular del virus de la gripe: *Hsw*1 *N*1, el virus de la gripe porcina. Se temió una pandemia mayor. El 24 de marzo, el entonces presidente Ford anunció una campaña a escala nacional sin precedentes para vacunar a todos los norteamericanos contra la gripe porcina. El congreso asignó 135 millones de dólares para financiar el esfuerzo, coordinado por oficinas de sanidad de los Departamentos de Sanidad, Educación y Bienestar, y llevado a cabo por departamentos de sanidad de los distintos estados de la Confederación. Después de sucesivos retrasos originados por las dudas sobre la dosis apropiada para niños, por dificultades de producción y por una seria controversia sobre seguro contra riesgo, se aprobó finalmente la vacunación el 1 de octubre. En noviembre se registraron varios casos de efectos secundarios neurológicos muy graves. El programa de inmunización o vacuna general se suspendió el 16 de diciembre y no se reanudó. No llegó a producirse la temida epidemia. En febrero de 1977, el nuevo Secretario de Sanidad, Educación y Bienestar, Joseph Califano, encargó una revisión del así llamado "fiasco de la gripe porcina" a Richard E. Neustadt y a Harvey V. Fineberg. El informe de éstos es una reconstrucción meticulosa de la "toma de decisiones en una enfermedad con complicaciones resbaladizas" y un examen detallado de las dificultades administrativas y técnicas que surgen siempre que se da una interacción entre la ciencia y la política.

El informe deja claro que hubo verdaderos motivos para preocuparse cuando se identificó la gripe porcina en Fort Dix. Este subtipo particular de virus de la gripe había sido el agente de la pandemia de 1918, que ocasionó la muerte a 20 millones de personas en todo el mundo y a 500.000 en los Estados Unidos. Sin embargo, desde finales de la década de 1920, la cepa se autoconfinó a los cerdos; ningún ser humano menor de 50 años podía haber creado anticuerpos contra dicho virus. La "deriva antigénica", ahora observada, significaba que lo que podía (o no) ser un virus de la gripe humana virulento había adquirido una nueva capa envolvente exterior de proteínas antigénicas que podían (o no) hacerlo muy contagioso para los humanos. Los directivos del Centro para Control de Enfermedades veían clara la posibilidad de una epidemia. Recomendaron hacer un serio esfuerzo no sólo para producir la vacuna contra la nueva cepa sino también para administrarla. Esta decisión original desembocó inexorablemente en la decisión del presidente Ford del mes de marzo. La decisión, concluyen Neustadt y Fineberg, no fue una cuestión de "política de partido" sino un juicio normativo válido para la salud pública. Sin embargo, y atendiendo a los beneficios a posteriori, se deduce que no se sopesó una opción mejor, a saber: almacenar una vacuna y

esperar a que el virus se extendiese de manera significativa antes de proceder a una vacunación general.

Los autores continúan analizando detenidamente cada paso de la toma de decisiones y ejecución de las medidas. Examinan con particular detalle el complicado asunto del riesgo, que surgió cuando las compañías de seguros rehusaron cubrir a los fabricantes de la vacuna contra el vasto peligro potencial de una campaña de inmunización general. Algunos funcionarios gubernamentales del Departamento de Sanidad previeron la contingencia ; debería haberse tratado de resolver al principio. Tal como estaba, el programa se retrasó; y el Congreso, por su parte, aprobó una nueva ley que especificaba que cualquier reclamación con respecto a la gripe porcina debería hacerse contra el gobierno, no contra los laboratorios.

La vacunación comenzó el 1 de octubre. Hubo mucha publicidad; la gente conocía el programa y la mayoría estaban dispuestos a secundarlo; durante los primeros 10 días se vacunaron más de un millón de personas. Después, primero en Pittsburgh y luego en otras partes, se registraron muertes subsiguientes a la vacuna. Las muertes se atribuyeron a cardiopatías y a otras causas coincidentes, pero algunos estados suspendieron la inmunización y el público se echó atrás.

El golpe final llegó en diciembre. Una rara enfermedad llamada síndrome de Guillain - Barré que producía parálisis se registró en varias personas que habían sido vacunadas. Todavía no está claro si el síndrome es un efecto secundario de la vacunación contra la gripe porcina, de cualquier otra vacuna, o posiblemente no tenga nada que ver con ella. Sin embargo, se puso inmediatamente de manifiesto que el programa tenía que suspenderse, aunque sólo fuera para determinar la asociación estadística del síndrome con la vacuna, de manera que el formulario de consentimiento que firmaban las personas que se vacunaban pudiese incluir alguna valoración del riesgo.

Neustadt y Fineberg no culpan a nadie en concreto. ni siquiera sacan las "moralejas", muchas de las cuales –piensan ellos– "surgen del relato". Se limitan a hacer "reflexiones" sobre cinco fenómenos que inciden en la toma de decisiones por parte del Secretario de Sanidad: "revisión de programas, análisis de la ejecución, reacciones de los medios, reputación de la oficina y enfermedades con implicaciones resbaladizas."

Las decisiones complicadas basadas en datos limitados requieren una reevaluación periódica. La necesaria revisión paso por paso exige, a su vez, que en cada etapa haya "un esbozo de las relaciones entre un punto muerto y una decisión" y "una exposición explícita de las suposiciones que subyacen a cada decisión". Por ejemplo, podía haber sido necesario decidir inmediatamente en marzo de 1976 si se debía o no producir la vacuna; no urgía tanto decidir si había que establecer un programa de inmunización general. Además, las suposiciones deben definirse y airearse, si es que han de revisarse. A uno le gustaría ver cada tema "planteado en todas sus partes componentes y razonado en términos probabilísticos".

En cuanto a la ejecución, "no es solamente algo que debe hacerse después de la decisión; se trata de algo que debe pensarse tanto o más que la decisión y antes que ésta, con el quid del problema". De esto hubo poco en el caso de la gripe porcina. Los autores creen que muchos de los problemas de la ejecución del programa se podían haber previsto. Llaman la atención sobre el valor de sopesar, "en la decisión, algunas estimaciones sobre las dificultades, probabilidades y consecuencias de equivocarse."

No se prestó suficiente atención a las realidades de la cobertura de las noticias: lo que los autores entienden no por "cobertura de influencia" sino por anticipación a ésta, preparación y balance de beneficios y costes. En un programa general esto es crucial a la hora de tomar medidas. "No se previó que los expertos que no estaban de acuerdo con el plan de vacunación general abrirían el fuego del debate" que da sabor a las noticias de televisión. No se preparó al público para incidentes tan predecibles como las muertes de Pittsburgh. No se hizo plan alguno para reaccionar ante la aparición de un efecto secundario como el del síndrome de Guillain - Barré.

La gripe es una "enfermedad resbaladiza", en primer lugar porque su expansión y periodo de aparición "estan hipotecados en el proceso de cambio antigénico" [véase "La epidemiología de la gripe", por Martin M. Kaplan y Robert G. Webster; INVESTIGACION Y CIENCIA, febrero de 1978]. La eficacia de una vacuna contra la gripe es de corta duración. Los síntomas de la gripe se han malinterpretado de forma generalizada, de manera que los profanos "y quizá la mitad de los doctores del país" aplican el término de forma incorrecta a varios trastornos gastrointestinales. El virus de la gripe puede incluso no ser la causa principal de los dolores y fiebre que se originan en una infección respiratoria. Por tanto, hasta una vacuna eficaz contra la gripe puede no parecerlo. Finalmente, es dificil estimar el impacto anual del virus de la gripe. Por todas estas razones, la gripe es mucho más "resbaladiza" que las enfermedades víricas combatidas hasta ahora con éxito mediante inmunización general, tales como la viruela, poliomielitis y paperas.

El telescopio de neutrinos

La astronomía de los neutrinos es una ciencia problemática. Puesto que se trata de una partícula sin carga eléctrica y sin masa aparente, un neutrino es capaz de atravesar el universo sin ser, no ya absorbido, sino ni siquiera desviado. Por tanto, deben ser visibles los neutrinos procedentes de las regiones del espacio donde la luz y las restantes radiaciones electromagnéticas estén obscurecidas. Sin embargo, por la misma razón, los neutrinos son sumamente difíciles de observar. Es esta misma tendencia a no interaccionar con la materia lo que da a la partícula su largo alcance; lo cual hace casi imposible capturar una de ellas en el laboratorio. En efecto, la mayor parte de los neutrinos que inciden sobre la Tierra la atraviesan fugazmente sin sufrir perturbación alguna, no dejando rastro a su paso.

A pesar de esta dificultad, durante los últimos tres años ha ido ganando apoyo un proyecto para construir un telescopio de neutrinos. Dicho proyecto recibe el nombre de DUMAND (del inglés "Deep Underwater Muon and Neutrino Detection", Detección Submarina Profunda de Muones y Neutrinos); el telescopio comprenderá un vasto sistema de detectores colocados cerca del fondo del océano. Un instrumento de este tipo puede servir tanto a los astrónomos, cuyo interés primario está en la fuente de los neutrinos, como a los físicos, que tendrían la posibilidad de estudiar las propiedades de los neutrinos a energías substancialmente mayores que aquellas que proporcionan los aceleradores de partículas.

Un rasgo característico de todo detector de neutrinos es su gran masa. Debido a que un neutrino presenta una probabilidad extremadamente pequeña de chocar con un cierto átomo, es necesario situar muchos átomos en su camino. Un experimento iniciado en la década de 1960 con el que recientemente se ha podido medir el flujo de neutrinos de baja energía emitidos por el Sol, empleaba como antena unas 675 toneladas del líquido de limpieza tetracloroetileno. Los neutrinos de alta energía, de interés astrofísico en el proyecto DUMAND, proceden de fuentes más distantes y llegan a la Tierra un número menor de ellos (al igual que la luz de las estrellas es más tenue que la luz solar). Se estima

que para la medida de este flujo menor se requiere un detector con una masa de unos mil millones de toneladas. El único material disponible en esta cantidad es el agua de mar. La idea básica consiste en distribuir sensores a través de un kilómetro cúbico del océano a una profundidad de alrededor de seis kilómetros.

El telescopio sería sensible sólo a los neutrinos de alta energía, pero incluso la gran mayoría de ellos pasarían a su través sin ser detectados. Los únicos neutrinos con una posibilidad de ser detectados serían aquellos pocos que interaccionasen con algún átomo en el agua, dando lugar a partículas cargadas eléctricamente que retrocederían con una alta energía. Estas partículas cargadas no serían capturadas, pero el aparato grabaría el trueno y el relámpago que marcan su paso. El "trueno" es un ruido de tono alto generado por la partícula cargada al ionizar átomos a lo largo de su trayectoria. El "relámpago" es la radiación de Cerenkov; se trata de la luz azul emitida por una partícula al moverse a través de un medio transparente a una velocidad mayor que la de la luz en aquel medio. (Estas velocidades son siempre menores que la de la luz en el vacío, velocidad que ninguna partícula puede sobrepasar.)

Para estas dos señales se han proyectado dos tipos de sensores. Las señales acústicas serían detectadas por unos cuantos grandes sistemas de hidrófonos, sistemas formados por unos 10.000 elementos cada uno. Se necesitarían sólo unos pocos sistemas de éstos porque el sonido posee un largo alcance en el agua, siendo limitado fundamentalmente sólo por la profundidad finita y la homogeneidad del océano. La radiación de Cerenkov se recogería, ampliando la señal, por tubos fotomultiplicadores. El alcance de la luz viene limitado por la absorción y la dispersión en el agua y, en consecuencia, depende de la claridad de ésta. Se ha observado que el agua oceánica profunda es en ciertas zonas más transparente que el agua destilada una sola vez; aún con todo, los detectores de luz probablemente deberían estar espaciados de manera uniforme por todo el volumen cada pocas decenas de metros. Parece que serán necesarios por lo menos 15.000 fotomultiplicadores, y quizá se llegue a los 125.000.

Los sensores acústicos y ópticos mandarían sus datos por cable a un computador central situado en la orilla. A partir de la información que éstos facilitasen, podría formarse una tosca imagen del rastro de una partícula. Además podrían deducirse la energía y la trayectoria original del neutrino incidente. Están en estudio dos emplazamientos para el sistema DUMAND: ambos se hallan en cuencas de subsidencia profundas cerca de las islas Hawai.

Existen varias posibles fuentes astrofísicas de neutrinos de alta energía. Pueden crearse neutrinos en la atmósfera exterior cuando los rayos cósmicos, principalmente protones, choquen con átomos del aire. Entre los productos de tales colisiones aparecen partículas de vida corta, en cuya desintegración se crean, entre otras cosas, neutrinos. De mayor interés astronómico son las posibles fuentes "puntuales", donde la presencia de partículas energéticas podría originar neutrinos de alta energía por un mecanismo de colisiones similar. Estos lugares incluyen supernovas, pulsars, quasars y los núcleos de varias galaxias activas. El telescopio podría resolver estas fuentes con una precisión superior al medio grado. Además, como los neutrinos penetran tanto la tierra como la atmósfera, el telescopio escrutaría continuamente la totalidad de la esfera celeste.

El núcleo de nuestra propia galaxia podría emitir neutrinos energéticos también. Se trata de una región cuya observación está prohibida a la astronomía óptica debido a la presencia de densas nubes de gas y polvo. Otra posibilidad es que los neutrinos de muy alta energía podrían haber sido creados en los primeros años del universo, poco después de la "gran explosión". En la actualidad su energía efectiva se habría reducido a causa de la expansión del universo, al igual que la longitud de onda de la luz se desplaza hacia el extremo rojo, o de menor energía, del espectro. Tales neutrinos cósmicos de fondo procederían de aquellas regiones del universo que son más remotas en el tiempo y en el espacio del sistema solar.

Se cree que el espectro de energías del flujo natural de neutrinos se extiende mucho más allá de las más altas energías disponibles en los actuales aceleradores de partículas, y de las de la totalidad de los aceleradores que se espera construir en los proximos años. Las energías de los aceleradores se están aproximando hoy en día a los 10^{12} electronvolt o un teraelectronvolt. DUMAND debería ser sensible a neutrinos cuyas energías sean superiores a un teraelectronvolt. Se presume que la proporción de sucesos inducidos por los neutrinos descenderá considerablemente al aumentar la energía, pero los cálculos indican que un kilómetro cúbico de agua de mar debería producir por lo menos unas cuantas interacciones cada día con energías superiores a los 10 teraelectronvolt, y algunas cada año con energías cercanas a los 100 teraelectronvolt. Estas cantidades se basan en el flujo mínimo estimado de neutrinos producidos en la atmósfera terrestre. Los neutrinos procedentes de otras fuentes podrían aumentar este flujo de forma substancial.

Los sucesos a estas energías pueden originar fenómenos nuevos en la física de partículas elementales. Uno que se ha esperado con impaciencia es la creación de bosones vectoriales intermedios, las partículas denominadas W^+, W^- *y* Z^0, que sirven como vehículos de una de las cuatro fuerzas básicas de la naturaleza, la fuerza débil. Precisamente porque los neutrinos sólo responden a la fuerza débil es por lo que son tan poco influenciados por la materia. También se deberán cruzar umbrales para la producción de otras partículas exóticas, como las partículas llamadas bosones de Higgs. Si existen otras especies adicionales de neutrinos, aparte de las dos conocidas hoy, bien pudieran observarse en los experimentos del DUMAND e incluso podría descrubrirse un cambio en el modelo de la misma fuerza débil. A las energías accesibles en la actualidad, la fuerza débil aumenta al crecer la energía de interacción, pero se espera que esta tendencia se estabilice.

La idea de emplear el océano como un detector de neutrinos fue concebida independientemente por varios científicos. El mismo proyecto DUMAND se estableció por un reducido grupo de investigadores, del que eran principales animadores Frederick Reines, de la Universidad de California en Irvine, John G. Learned, que está en Irvine con excedencia de la Universidad de Wisconsin, y Arthur Roberts, del Fermi Nacional Accelerator Laboratory. Reines, que es el presidente del comité que dirige el proyecto, colaboró con Clyde L. Cowan, Jr., en la primera detección experimental de neutrinos, en 1956. En la actualidad están afiliados a la organización DUMAND unos cien investigadores.

Han finalizado ya algunos estudios preliminares de los emplazamientos propuestos y de la física de los detectores acústicos y ópticos. Se espera que, en los primeros meses de este año, se inicie un proyecto piloto destinado a investigar la tecnología de los detectores y a demostrar la viabilidad del proyecto global. Una meta importante del programa piloto es establecer el costo de la instalación a gran escala. Reines estima que el costo no puede bajar de algunas decenas de millones de dólares, pero depende en gran manera de factores tales como el número de elementos del sistema de sensores ópticos. La construcción y la instalación requerirán probablemente de cinco a diez años.

La estructura de la hemoglobina y el transporte respiratorio

La hemoglobina transporta oxígeno desde los pulmones a los tejidos y colabora en el transporte del dióxido de carbono devolviéndolo hacia los pulmones. Este doble papel lo lleva a cabo transformándose en dos estructuras alternativas

M. F. Perutz

¿Por qué la hierba es verde, o por qué
nuestra sangre es roja?
Son misterios que nadie ha alcanzado
a descifrar.
En este estadio inferior, pobre alma,
¿qué harás?

JOHN DONNE
"Del Progreso del Alma"

Era yo estudiante, cuando andaba tras la solución de un gran problema de bioquímica. Un día salí de Viena, mi ciudad natal, en busca del Gran Sabio en Cambridge. El me enseñó que el enigma de la vida estaba escondido en la estructura de las proteínas, y que la cristalografía por rayos X era el único método capaz de resolverlo. El Sabio era John Desmond Bernal, quien acababa de descubrir los ricos perfiles de difracción de rayos X proporcionados por las proteínas cristalinas. Nosotros le llamábamos el "Sabio", efectivamente, porque lo sabía todo. Me convertí en su discípulo.

En 1937 escogí la hemoglobina como proteína cuya estructura quería resolver, pero la estructura demostró ser mucho más compleja que cualquiera de las resueltas con anterioridad, y así tardé más de 20 años en encontrar la solución. La promesa del Sabio se cumplió en 1959, cuando Ann F. Cullis, Hilary Muirhead, Michael G. Rossmann, Tony C. T. North y yo mismo desentrañamos, por primera vez, el perfil arquitectónico de la molécula de hemoglobina [véase "The Hemoglobin Molecule", por M. F. Perutz; SCIENTIFIC AMERICAN, noviembre, 1964]. Nos sentimos como exploradores que han descubierto un nuevo continente; pero aquello no era el final del viaje, puesto que nuestro muy admirado modelo no revelaba su maquinaria interna: no nos proporcionó ninguna pista acerca de los mecanismos moleculares del transporte respiratorio. ¿Por qué no? Los colegas bien intencionados sugirieron en seguida que nuestra estructura, laboriosamente conseguida, no era más que un artefacto de la cristalización y podría diferir bastante de la estructura de la hemoglobina en su medio ambiente natural, que es el glóbulo rojo de la sangre.

La hemoglobina es la proteína vital que transporta oxígeno desde los pulmones a los tejidos y facilita la vuelta del dióxido de carbono desde los tejidos a los pulmones. Estas funciones y su sutil interacción también contribuyen a que la hemoglobina sea una de las proteínas más interesantes de estudiar. Como todas las proteínas, consta de pequeñas moléculas orgánicas denominadas aminoácidos, unidos unos a otros en una secuencia lineal llamada cadena polipeptídica. Los aminoácidos son de 20 clases diferentes y su secuencia en la cadena está determinada genéticamente. Una molécula de hemoglobina está compuesta por cuatro cadenas polipeptídicas, dos cadenas alfa de 141 restos de aminoácidos cada una y dos cadenas beta de 146 restos cada una. Las cadenas alfa y beta tienen diferentes secuencias de aminoácidos, pero se entremezclan para formar estructuras tridimensionales similares. Cada cadena contiene un grupo hemo, que le da a la sangre su color rojo. El grupo hemo consiste en un anillo de átomos de carbono, nitrógeno e hidrógeno, denominado porfirina, con un átomo de hierro, a modo de gema, en el centro. Una única cadena polipeptídica combinada con un solo hemo se llama subunidad de hemoglobina o monómero de la molécula. En la molécula completa las cuatro subunidades están estrechamente entrelazadas, como un rompecabezas tridimensional, para formar un tetrámero.

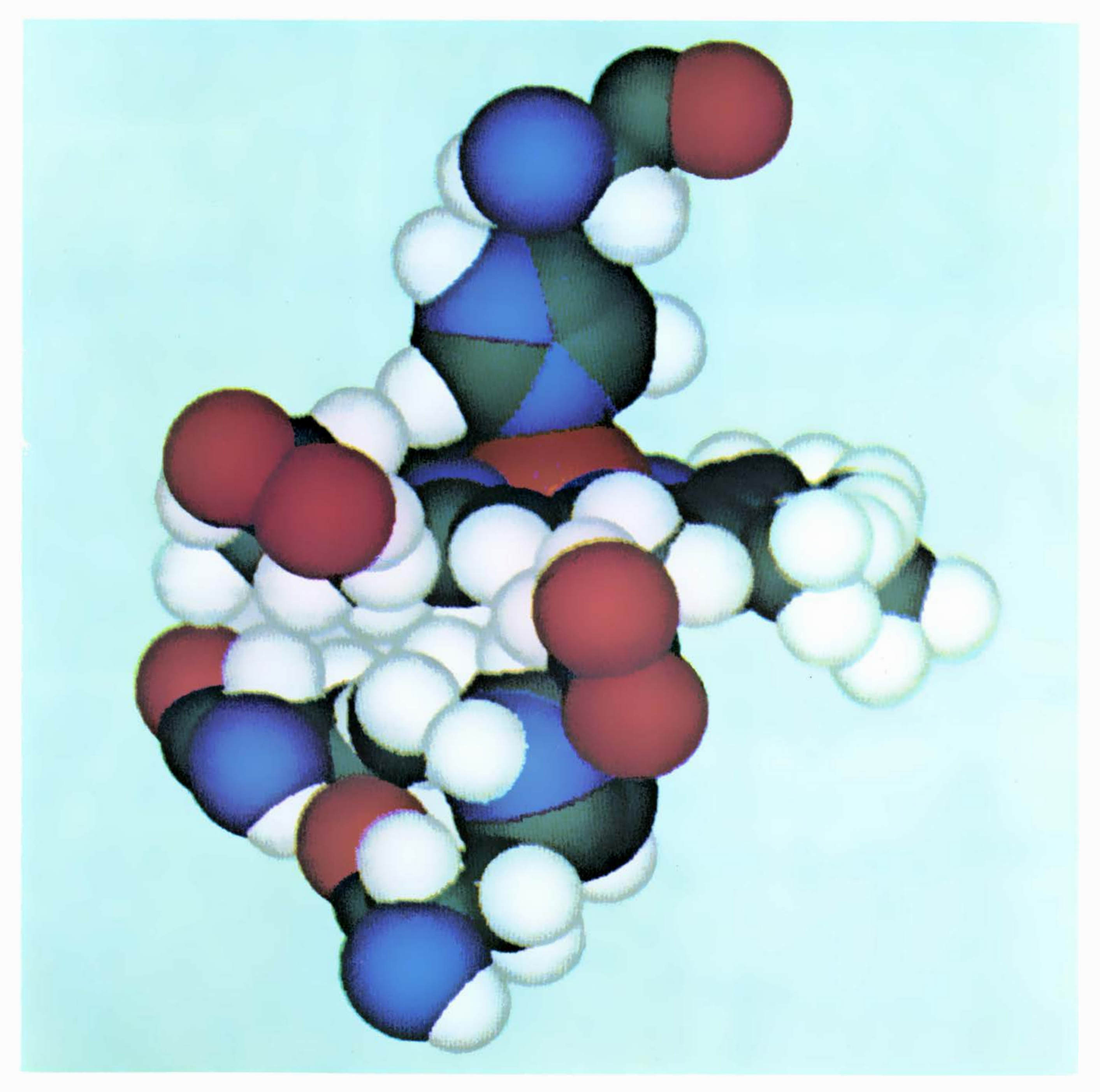
En el músculo rojo hay otra proteína, denominada mioglobina, de constitución y estructura similar a la subunidad beta de la hemoglobina, pero formada solamente por una cadena polipeptídica y un hemo. La mioglobina se combina con el oxígeno liberado por los glóbulos rojos, y lo almacena y lo transporta a unos orgánulos intracelulares denominados mitocondrias, donde el oxígeno genera energía química gracias a la combustión de glucosa en dióxido de carbono y agua. La mioglobina fue la primera proteína cuya estructura tridimensional se determinó; dicha estructura la resolvieron mi colega John C. Kendrew y sus colaboradores.

La función de la hemoglobina

La mioglobina es la más sencilla de ambas moléculas. Esta proteína, con sus 2500 átomos de carbono, nitrógeno, oxígeno, hidrógeno y azufre, existe con el único propósito de permitir que su exclusivo átomo de hierro forme un enlace químico débil con la molécula de oxígeno (O_2). ¿Por qué la naturaleza se toma tanto trabajo para llevar a cabo lo que, en apariencia, es algo tan simple? Como la mayoría de los compuestos de hierro, el grupo hemo se combina por sí mismo con el oxígeno tan fuertemente que el enlace, una vez formado, es difícil de romper. Esto ocurre porque un átomo de hierro puede existir en dos estados de valencia: hierro ferroso, que transporta dos cargas positivas, como en el sulfato de hierro, que es el que se recomienda a las personas anémicas, y el hierro férrico, que transporta tres cargas positivas, como el óxido de hierro o la herrumbre. Normalmente, el hemo ferroso reacciona con el oxígeno de modo irreversible para formar hemo férrico, pero cuando el hemo ferroso está aco-

plado en los pliegues de la cadena de globina, queda protegido de suerte que la reacción con el oxígeno es reversible. El efecto de la globina en la química del hemo ha sido explicado hace tan sólo poco tiempo, con el descubrimiento de que la oxidación irreversible del hemo es consecuencia de un compuesto intermedio con el que la molécula de oxígeno forma un puente entre los átomos de hierro de los dos hemos. En la mioglobina y la hemoglobina, los pliegues de la cadena polipeptídica evitan la formación de dicho puente mediante el aislamiento de cada hemo en cavidades separadas. Además, en la proteína el hierro está unido a un átomo de nitrógeno del aminoácido histidina, el cual cede cargas negativas que capacitan al hierro para formar un enlace débil con el oxígeno.

Una solución, sin oxígeno, de mioglobina o hemoglobina es de color púrpura como la sangre venosa; cuando se bombea oxígeno a través de dicha solución, ésta se torna escarlata como la sangre arterial. Si estas proteínas han de actuar como transportadores de oxígeno, la hemoglobina deberá poder captar entonces el oxígeno en los pulmones, donde está en abundancia, y cederlo a la mioglo-



GRUPO HEMO, centro activo de la molécula de la hemoglobina; constituye el lugar de enlace del oxígeno. El hemo es un anillo plano, denominado porfirina, con un átomo de hierro en su centro; se observa aquí el borde extendiéndose horizontalmente por la parte media de la ilustración. Se muestran también tres de los 16 restos de aminoácidos de la globina que están en contacto con el hemo. En esta imagen, generada por ordenador, cada átomo está representado por una esfera en la que no pueden penetrar otros átomos a menos que los átomos estén enlazados químicamente; donde dos átomos están enlazados, las esferas se ocluyen en parte. Los átomos de carbono aparecen en negro, los de nitrógeno en azul, los de oxígeno en rojo, los átomos de hidrógeno en blanco y el átomo de hierro de coloración oxidada. El modelo muestra el hemo desoxigenado; el oxígeno se enlaza en el lado inferior del átomo de hierro. La imagen fue obtenida por Richard J. Feldmann y Thomas K. Porter, de los National Institutes of Health, a partir de las coordenadas atómicas determinadas por Giulio Fermi, del Consejo de I. Médica del Laboratorio de Biología Molecular de Cambridge.

bina en los capilares musculares, donde es menos abundante; la mioglobina pasa, a su vez, el oxígeno a las mitocondrias, donde éste escasea más todavía.

Un simple experimento muestra que la mioglobina y la hemoglobina pueden llevar a cabo este intercambio, habida cuenta de la existencia de un equilibrio entre el oxígeno libre y el oxígeno ligado al hierro del hemo. Supóngase que se coloca una solución de mioglobina en un recipiente construido de tal manera que pueda mezclarse un gran volumen de gas con esta solución y que el color pueda medirse por medio de un espectroscopio. Sin oxígeno solamente se observa el color púrpura de la desoximioglobina. Si se inyecta un poco de oxígeno, parte del oxígeno se combina con parte de la desoximioglobina para formar oximioglobina, que es escarlata. El espectroscopio mide la proporción de oximioglobina en la solución. El suministro de oxígeno y las mediciones del espectroscopio se repiten hasta que toda la mioglobina se ha tornado escarlata. Los resultados se representan en una gráfica con la presión parcial de oxígeno en el eje horizontal y el porcentaje de la oximioglobina en el eje vertical. La gráfica tiene la forma de una hipérbola rectangular: empinada al principio, cuando todas las moléculas de mioglobina están libres; aplanada al final, cuando las moléculas libres de mioglobina son tan escasas que solamente presiones elevadas de oxígeno las pueden saturar.

Para comprender este equilibrio hay que contemplar su dinámica. Bajo la influencia del calor, las moléculas de la solución y del gas se mueven rápida e irregularmente y están en constante colisión. Las moléculas de oxígeno que entran y salen de la solución van formando y rompiendo enlaces con las moléculas de la mioglobina. El número de enlaces hierro-oxígeno que se rompen en un segundo es proporcional al número de moléculas de oximioglobina. El número de enlaces que se forman en un segundo es proporcional a la frecuencia de colisiones entre la mioglobina y el oxígeno, la cual a su vez viene determinada por el producto de sus concentraciones. Cuanto más oxígeno se añade al gas, más moléculas de oxígeno se disuelven, colisionan y se enlazan con la mioglobina; esto aumenta el número de moléculas de oximioglobina presentes, y, por consiguiente, también el número de enlaces hierro-oxígeno posibles de quebrar, hasta que el número de moléculas de mioglobina combinadas con el oxígeno en un segundo se vuelve igual al número de las que pierden su oxígeno en un segundo. Cuando esto ocurre, es que ha quedado establecido un equilibrio químico.

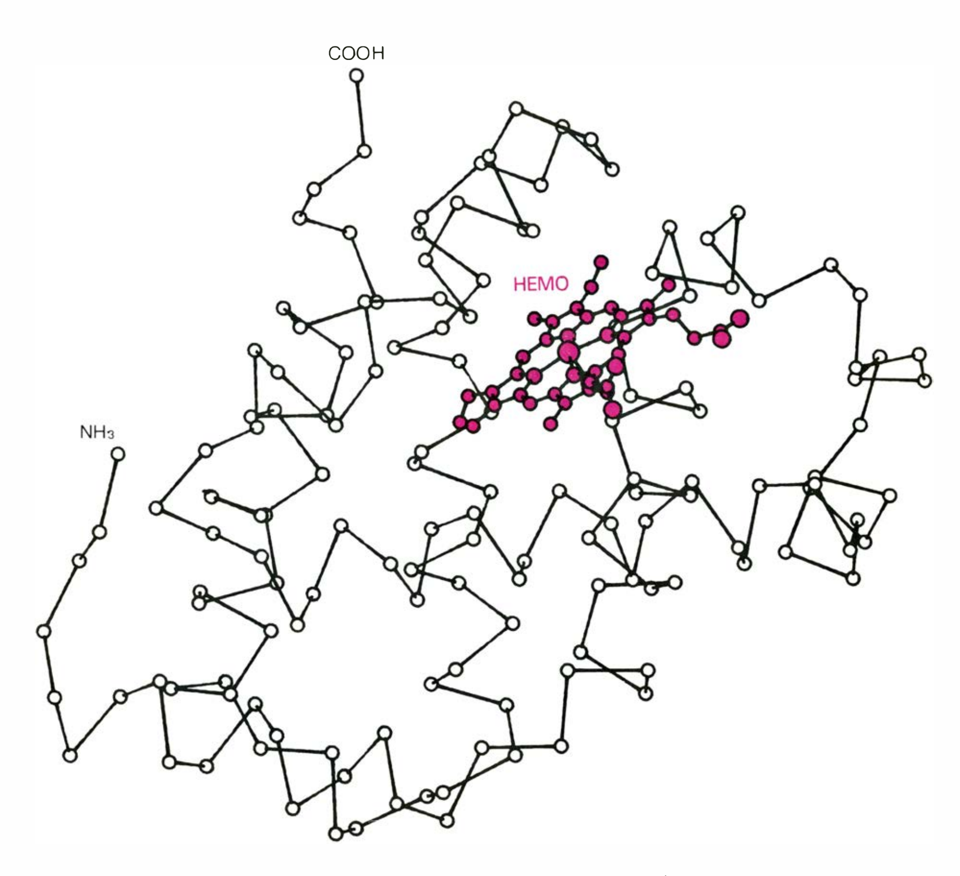
El equilibrio queda mejor determinado por una gráfica en la que el logaritmo del cociente entre las moléculas de oximioglobina (Y) y las moléculas de desoximioglobina ($1-Y$) se representa en función del logaritmo de la presión parcial de oxígeno. La hipérbola se convierte ahora en una línea recta a 45 grados de los ejes. La intersección de la línea con el eje horizontal que se produce cuando $Y/(1-Y) = 1$ proporciona la constante de equilibrio K. Esta es la presión parcial de oxígeno a la que la mitad exacta de las moléculas de mioglobina han captado oxígeno. Cuanta mayor es la afinidad de la proteína por el oxígeno, menor es la presión necesaria para alcanzar la saturación media y menor es la constante de equilibrio. La pendiente de 45 grados permanece constante, aunque una más baja afinidad hacia el oxígeno desplace la línea hacia la derecha y una afinidad más alta desplace a ésta hacia la izquierda.

Si el mismo experimento se hace con sangre o con una solución de hemoglobina, se obtiene un resultado totalmente diferente. La curva asciende, al principio, ligeramente; luego se vuelve empinada y, por último, se aplana aproximándose a la curva de mioglobina. Esta extraña forma sigmoidea significa que las moléculas carentes de oxígeno (desoxihemoglobina) son reacias a captar las primeras moléculas de oxígeno, aunque su apetito por el oxígeno aumenta a medida que éste es satisfecho. A la inversa, la pérdida de oxígeno por algunos de los hemos disminuye la afinidad del oxígeno de los restantes. La distribución de oxígeno entre las moléculas de hemoglobina en una solución sigue, por consiguiente, la parábola bíblica del rico y del pobre: "Los que ya tienen algo recibirán más y se enriquecerán, y al que tenga poco se le quitará lo poco que tenga". Este fenómeno sugiere que hay un cierto tipo de comunicación entre los

ESTRUCTURA QUIMICA del grupo hemo y de los aminoácidos circundantes. La estructura química se muestra por medio de un esqueleto de líneas que conectan los centros de los átomos. El único enlace químico entre el grupo hemo y la proteína que engarza a éste es el enlace entre el átomo de hierro y el aminoácido de la parte superior, denominado histidina proximal; los dos aminoácidos de la parte inferior (la histidina distal y la valina distal) alcanzan al hemo, pero no se hallan enlazados a él. La histidina proximal constituye la principal vía de comunicación entre el grupo hemo y el resto de la molécula. En el estado desoxigenado que aquí aparece, el hierro sobresale por encima de la porfirina y su vuelta a una posición centrada puede ser impedida mediante la repulsión entre un ángulo de la histidina proximal y uno de los átomos de nitrógeno de la porfirina. Clave construida por R. Diamond.

hemos de cada molécula; ello explica por qué los fisiólogos la han denominado interacción hemo-hemo.

El mecanismo de la interacción hemo-hemo se comprenderá mejor si trazamos una gráfica de representación logarítmica. La curva de equilibrio se inicia con una línea recta a 45 grados de los ejes, pues las primeras moléculas de oxígeno son tan escasas que sólo un hemo de cada molécula tiene la oportunidad de atrapar a uno de ellos, y todos los hemos, por consiguiente, reaccionan de una forma independiente, como en el caso de la mioglobina. A medida que llega más oxígeno, los cuatro hemos de cada molécula empiezan a interaccionar, y la curva se hace más empinada. La tangente a su máxima pendiente se conoce como coeficiente de Hill (n), en honor del fisiólogo A. V. Hill, quien fue el primero en realizar un análisis matemático del equilibrio del oxígeno. El valor normal del coeficiente de Hill oscila en torno a 3; sin la interacción hemo-hemo se convierte en la unidad. La curva finaliza con otra línea a 45 grados de los ejes, lo cual explica por qué el oxígeno se ha vuelto ahora tan abundante que sólo el último hemo de cada molécula tiene la probabilidad de estar libre, y todos los hemos en la solución reaccionan independientemente, una vez más.



LA SUBUNIDAD DE HEMOGLOBINA consta de un grupo hemo (*color*) envuelto en una cadena polipeptídica. El polipéptido es una secuencia lineal de restos de aminoácidos, cada uno de los cuales está representado aquí por un simple círculo, que señala la posición del átomo de carbono central (alfa). La cadena empieza con un grupo amino (NH_3) y termina con un grupo carboxilo (COOH). La mayor parte del polipétido está arrollado para formar segmentos helicoidales, pero también hay regiones no helicoidales. El diagrama, generado por un ordenador, de una subunidad de hemoglobina de caballo fue preparado por Feldmann y Porter. En la figura anterior se recurrió también al ordenador.

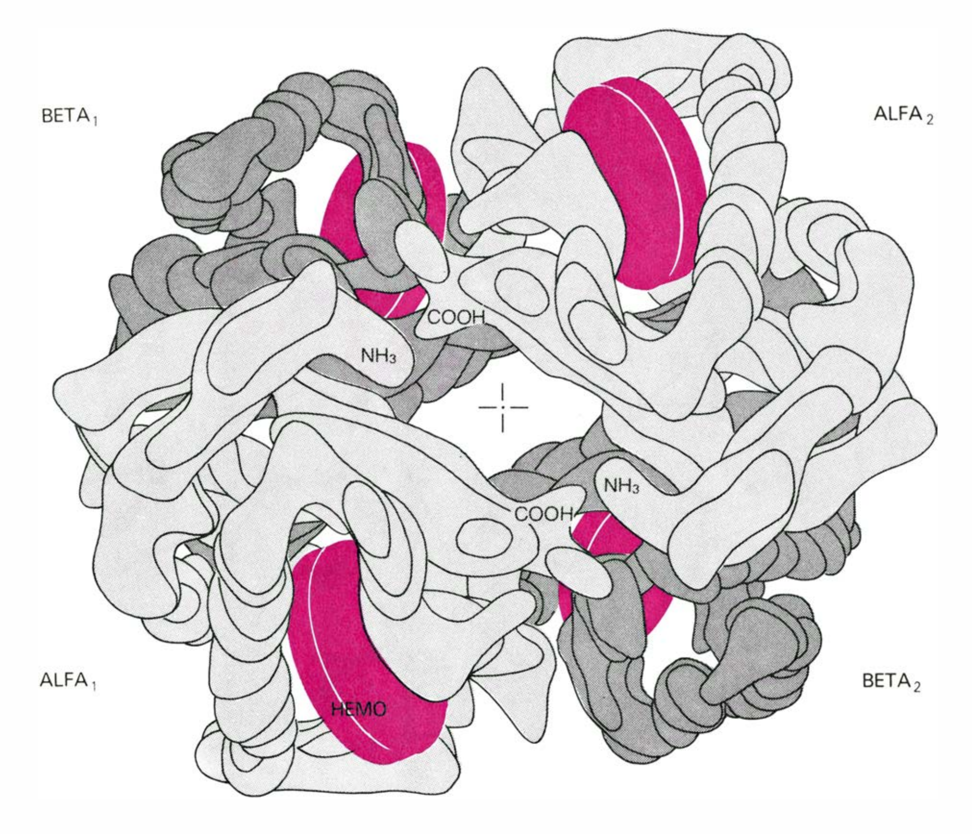
Efectos de cooperatividad

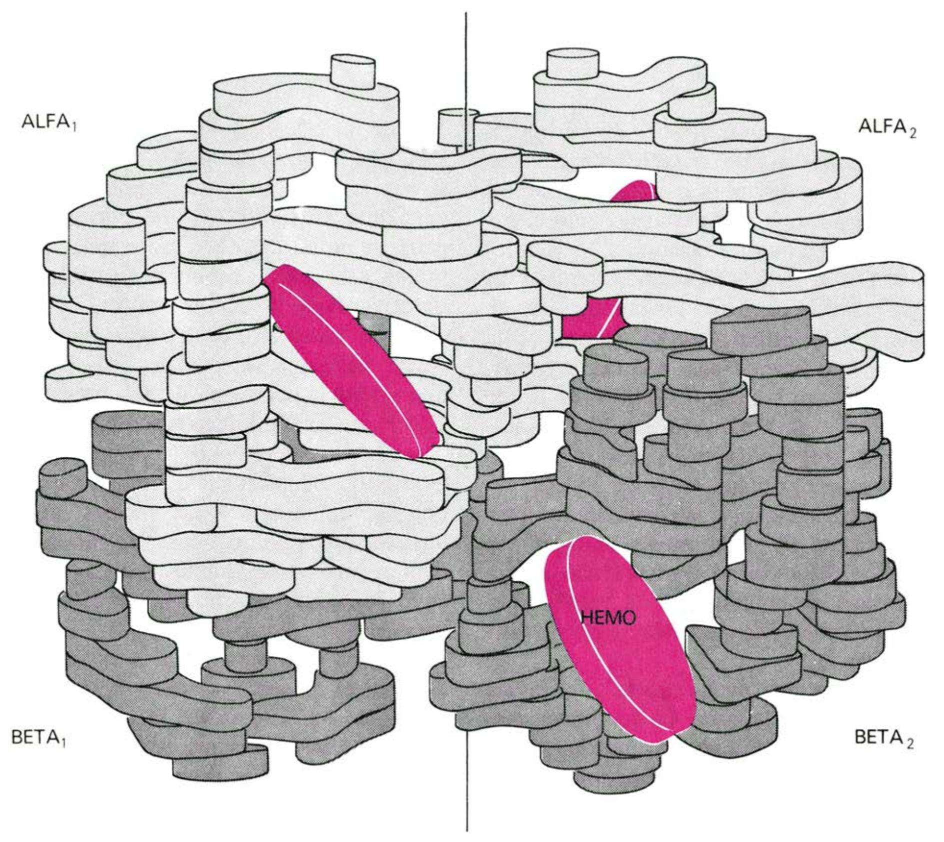
El coeficiente de Hill y la afinidad del oxígeno por parte de la hemoglobina dependen de la concentración de diversos factores químicos en el glóbulo rojo sanguíneo: protones (átomos de hidrógeno sin electrones, cuya concentración puede medirse como pH), dióxido de carbono (CO_2), iones cloruro (Cl^-) y un compuesto de ácido glicérico y fosfato denominado 2,3-difosfoglicerato (DPG). Al aumentar la concentración de cualquiera de estos factores, la curva de equilibrio del oxígeno se desplaza hacia la derecha, hacia una menor afinidad de oxígeno, haciendo la curva más sigmoidea. El aumento de la temperatura también desplaza la curva hacia la derecha, pero la hace menos sigmoidea. Lo extraño es que ninguno de estos factores, a excepción de la temperatura, influye en la curva de equilibrio del oxígeno de la mioglobina, a pesar de que la química y la estructura de la mioglobina están muy relacionadas con la química y la estructura de las distintas cadenas de la hemoglobina.

¿A qué obedecen estos efectos extraordinarios? ¿Por qué no le basta al glóbulo rojo poseer un simple transportador de oxígeno como es la mioglobina? Dicho transportador no permitiría que se descargara el suficiente oxígeno desde el glóbulo rojo a los tejidos, ni permitiría que fuera transportado el suficiente dióxido de carbono a los pulmones por el plasma sanguíneo. La presión parcial de oxígeno en los pulmones es de unos 100 milímetros de mercurio, la cual es suficiente para saturar la hemoglobina con oxígeno, sea la curva de equilibrio sigmoidea o sea hiperbólica. En la sangre venosa la presión sanguínea es aproximadamente de 35 milímetros de mercurio; si la curva fuera hiperbólica, se liberaría menos del 10 por ciento del oxígeno transportado a esta presión, de manera que un hombre podría asfixiarse aun cuando respirase normalmente.

Cuanto más pronunciada es la forma sigmoidea de la curva de equilibrio, mayor es la fracción de oxígeno que puede liberarse. Varios factores influyen en esta acción. La oxidación de las sustancias nutritivas por los tejidos libera ácido láctico y ácido carbónico; estos ácidos liberan, a su vez, protones, que desplazan la curva hacia la derecha, hacia una afinidad de oxígeno menor, y la hacen más sigmoidea. Otro importante regulador de la afinidad del oxígeno es el DPG. El número de moléculas de DPG que se encuentran en el glóbulo rojo es aproximadamente el mismo que el número de moléculas de hemoglobina, 280 millones; y resulta probable que permanezca bastante constante durante la circulación. La escasez del oxígeno, sin embargo, hace que se produzca más DPG, que facilita la liberación de más oxígeno. Con la curva sigmoidea típica puede cederse casi la mitad del oxígeno transportado a los tejidos. El feto humano tiene una hemoglobina con las mismas cadenas alfa que la hemoglobina del humano adulto, pero con diferentes cadenas beta, cuyo resultado es una menor afinidad hacia el DPG. Esto determina que la hemoglobina fetal tenga una mayor afinidad por el oxígeno y facilita la transferencia de oxígeno desde la circulación materna a la circulación fetal.

El monóxido de carbono (CO) se combina con el hierro del hemo en el mismo lugar que el oxígeno, pero su afinidad es 150 veces mayor; el monóxido de carbono desplaza, por tanto, al oxígeno, y ello explica su gran toxicidad. En fumadores empedernidos hasta el 20 por ciento de los lugares de combinación del oxígeno





MOLECULA COMPLETA de hemoglobina. Consta de cuatro subunidades, cada una de las cuales está formada por una cadena polipeptídica y un grupo hemo. Hay dos clases de subunidades, alfa *(blanco)* y beta *(gris)*, que tienen diferentes secuencias de restos de aminoácidos, pero similar estructura tridimensional. La cadena beta tiene una hélice adicional. Las cuatro subunidades, que se observan aquí en dos planos, están dispuestas en los vértices de un tetraedro alrededor de un eje de simetría doble. Cada grupo hemo *(color)* se dispone en una cavidad separada en la superficie de la molécula.

pueden estar bloqueados por monóxido de carbono, de forma que la sangre transporta menos oxígeno. Además, el monóxido de carbono tiene otro efecto más funesto todavía. La combinación de uno de los cuatro hemos de cualquier molécula de hemoglobina con el monóxido de carbono aumenta la afinidad hacia el oxígeno de los tres hemos restantes en razón de la interacción hemo-hemo. La curva de equilibrio del oxígeno se desplaza, por tanto, hacia la izquierda, con lo que disminuye la fracción de oxígeno transportado que puede cederse a los tejidos.

Si los protones menguan la afinidad de la hemoglobina por el oxígeno, entonces las leyes de acción y reacción exigen que el oxígeno disminuya la afinidad de la hemoglobina por los protones. La liberación del oxígeno causa que la hemoglobina se combine con los protones y viceversa; suelen captarse dos protones por cada cuatro moléculas de oxígeno liberado, y, a su vez, se liberan dos protones cuando se captan cuatro moléculas de oxígeno. A esta acción recíproca se la conoce como efecto Bohr y es la clave del mecanismo del transporte de dióxido de carbono. El dióxido de carbono liberado por los tejidos en la respiración es demasiado insoluble para ser transportado como tal, pero se puede volver más soluble combinándolo con agua para formar un ion bicarbonato y un protón. La reacción química se escribe así

$$CO_2 + H_2O \rightarrow HCO_3^- + H^+.$$

En ausencia de hemoglobina esta reacción se pararía pronto debido al exceso de protones producidos, de la misma manera que un fuego se extingue cuando la chimenea está bloqueada. La desoxihemoglobina actúa como un tampón, captando los protones y desplazando el equilibrio hacia la formación de bicarbonato soluble. En los pulmones ocurre el proceso inverso. Allí, al unirse el oxígeno a la hemoglobina, los protones son expulsados, conduciendo al dióxido de carbono fuera de la solución de suerte que éste pueda ser espirado. La reacción entre el dióxido de carbono y el agua está catalizada por la anhidrasa carbónica, un enzima de los glóbulos rojos. El enzima acelera la reacción hasta una velocidad aproximada de medio millón de moléculas por segundo, una de las más rápidas de todas las reacciones biológicas conocidas.

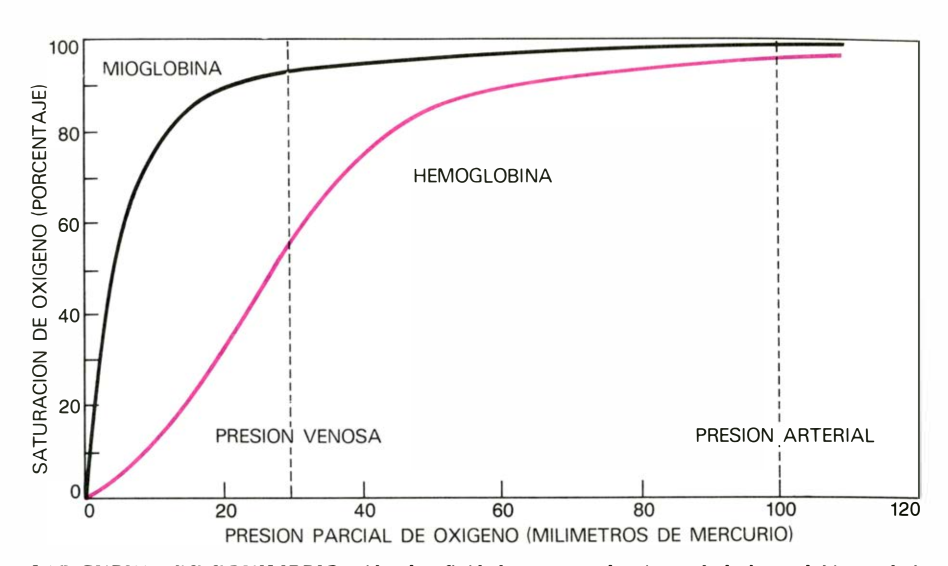
Aunque menos importante, existe un segundo mecanismo para el transporte de dióxido de carbono. Este gas se enlaza con mayor facilidad a la desoxihemoglobina que a la oxihemoglobina, de manera que tiende a ser captado cuando el oxígeno se libera y, expulsado, cuando el oxígeno es

atraído. Ambos mecanismos de transporte de dióxido de carbono son antagónicos: por cada molécula de dióxido de carbono unida a la desoxihemoglobina, se libera uno o dos protones, lo cual se opone a la conversión de otras moléculas de dióxido de carbono en bicarbonato. Los protones cargados positivamente que entran en el glóbulo rojo arrastran iones cloruro con carga negativa a su paso; estos iones se unen más fácilmente también a la desoxihemoglobina que a la oxihemoglobina. El DPG se sintetiza en el mismo glóbulo rojo y no puede salir a través de la membrana celular. Está fuertemente unido a la desoxihemoglobina y sólo muy débilmente unido a la oxihemoglobina.

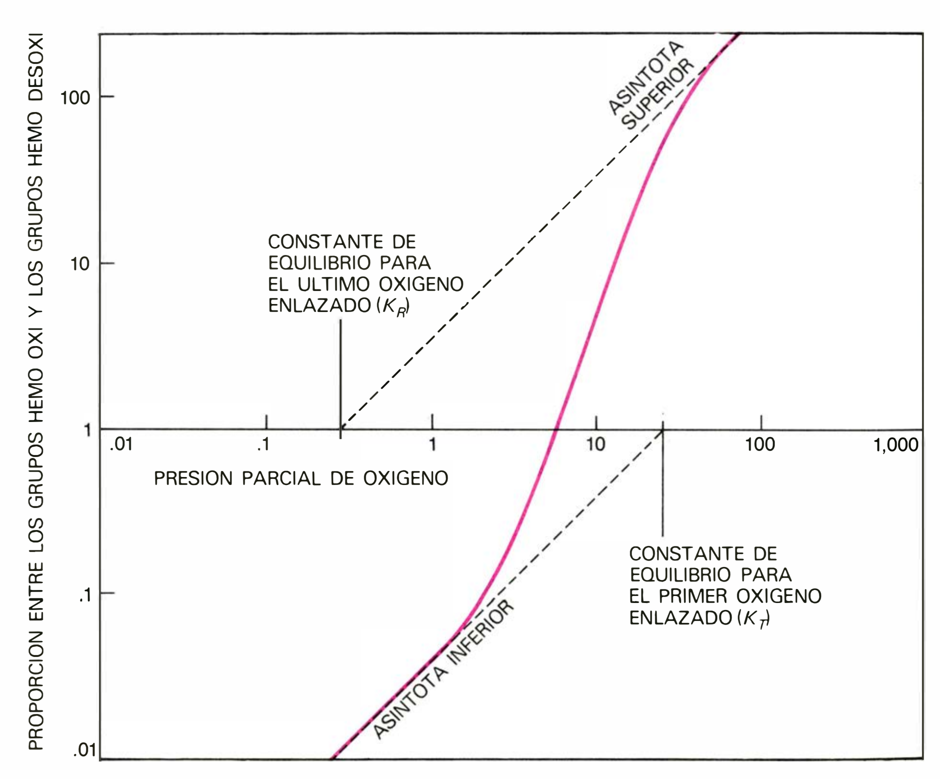
La interacción hemo-hemo y la relación entre el oxígeno y los otros cuatro ligandos se denominan, colectivamente, efectos de cooperatividad de la hemoglobina. El descubrimiento de los mismos por una sucesión de eminentes fisiólogos y bioquímicos duró más de medio siglo y trajo consigo muchas disputas. En 1938, Felix Haurowitz, de la Universidad Carolina de Praga, hizo otra observación vital. Descubrió que la desoxihemoglobina y la oxihemoglobina formaban cristales diferentes, como si fueran sustancias químicas diferentes; ello implicaba que la hemoglobina no era un tanque de oxígeno sino un pulmón molecular, ya que cambia su estructura cada vez que capta o libera oxígeno.

La teoría alostérica

El descubrimiento de la interacción entre los cuatro hemos dejó claro que éstos debían estar en contacto, pero en la ciencia lo que es obvio no resulta necesariamente cierto. Cuando se resolvió finalmente la estructura de la hemoglobina, se encontró que los hemos se hallaban aislados en cavidades por la superficie de las subunidades. Sin contacto entre sí, ¿cómo podía uno de ellos percibir si los otros se habían combinado con oxígeno? Y ¿cómo podía una serie tan heterogénea de agentes químicos, como los protones, iones cloruro, dióxido de carbono y difosfoglicerato, influir en la curva de equilibrio del oxígeno de manera parecida? No parecía plausible que cualquiera de ellos se llegara a unir directamente con los hemos, que todos ellos se pudieran unir en cualquier otro lugar común. Pero de nuevo resultó que andábamos equivocados. Para enredar el misterio, ninguno de estos agentes afectaba al equilibrio de oxígeno de la mioglobina o de las subunidades aisladas de la hemoglobina. Ahora sabemos que todos los efectos de cooperatividad desaparecen si la molécula de hemoglobina se divide simplemente por la mitad, pero esta



LAS CURVAS DE EQUILIBRIO miden la afinidad para con el oxígeno de la hemoglobina y de la molécula más simple, la mioglobina. La mioglobina, que es una proteína muscular, tiene solamente un grupo hemo y una cadena polipeptídica, y se parece a una única subunidad de hemoglobina. El eje vertical presenta la cantidad de oxígeno unido a una de estas proteínas, expresado en porcentaje de la cantidad total que puede ser unido. El eje horizontal mide la presión parcial de oxígeno en una mezcla de gases en la que se permite que la solución alcance el equilibrio. Para la mioglobina *(negro)*, la curva de equilibrio es hiperbólica. La mioglobina absorbe oxígeno rápidamente y alcanza la saturación a baja presión. La curva de la hemoglobina *(color)* es sigmoidea: al principio, la hemoglobina se muestra reticente a captar oxígeno, pero su afinidad aumenta a medida que va combinándose con él. A presión de oxígeno arterial, ambas moléculas están casi saturadas: a presión venosa, sin embargo, la mioglobina cedería sólo hasta un 10 por ciento de su oxígeno, mientras que la hemoglobina liberaría aproximadamente la mitad. A cualquier presión parcial, la mioglobina tiene una mayor afinidad que la hemoglobina, lo que permite que el oxígeno se transfiera desde la sangre hasta el músculo.



FORMA SIGMOIDEA de la curva de equilibrio del oxígeno, que aparece más pronunciada cuando la saturación fraccional y la presión parcial de oxígeno se representan en escalas logarítmicas. En dicha gráfica, la curva de equilibrio de la mioglobina se convierte en una línea recta a 45 grados de los ejes. La curva de la hemoglobina empieza y termina con líneas rectas, denominadas asíntotas, con el mismo ángulo. Sus intercepciones con la línea horizontal trazada donde las concentraciones de desoxihemoglobina y oxihemoglobina son iguales proporcionan las constantes de equilibrio para la primera y última molécula de oxígeno que se combina con la hemoglobina. En la interpretación alostérica de la curva, éstas son, respectivamente, la constante de equilibrio de la estructura *R* (K_R) y la constante de equilibrio de la estructura *T* (K_T). En la curva presentada, la primera constante vale 30 y 0,3, la segunda; ello indica que la afinidad para con el último oxígeno enlazado es 100 veces mayor que la afinidad para con el primero. Esta proporción determina la energía libre de la interacción hemo-hemo, que mide el grado de influencia ejercida por la combinación de cualquiera de los cuatro hemos con el oxígeno sobre la afinidad del oxígeno de los hemos restantes. Si el principio y final de la curva no pueden medirse con exactitud, la máxima pendiente (coeficiente de Hill), indica el grado de la interacción hemo-hemo.

MIOGLOBINA

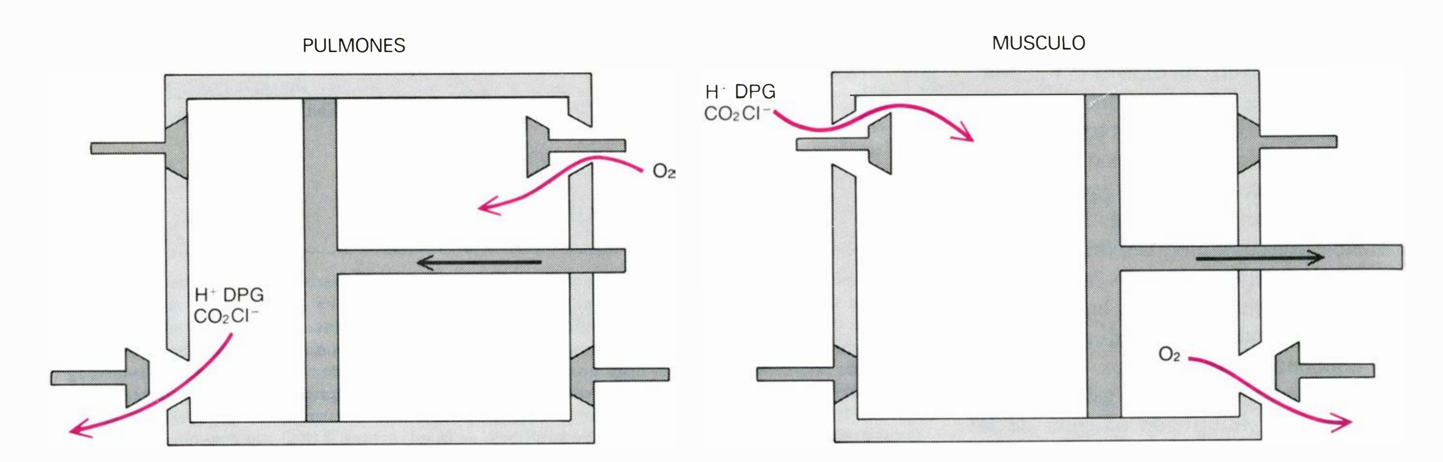
a b c d

HEMOGLOBINA

0 20 100 30

PRESION PARCIAL DE OXIGENO (MILIMETROS DE MERCURIO)

CON LA INTERACCION HEMO-HEMO se facilita la liberación de oxígeno por parte de la hemoglobina; lo que no se observa en la mioglobina. Las proteínas aparecen en gris en el estado desoxi y coloreadas cuando están oxigenadas; hay cuatro veces más moléculas de mioglobina que de hemoglobina, aunque igual número de lugares de enlace. En ausencia de oxígeno (*a*), están vacíos casi todos los sitios de enlace. A medida que la presión de oxígeno aumenta (*b*), la mioglobina capta el oxígeno más de prisa, pero cuando la presión alcanza la de los pulmones (*c*), ambos transportadores están próximos a la saturación. La distinción entre las dos moléculas se hace más aparente cuando la presión de oxígeno desciende al nivel típico del sistema capilar (*d*). El equilibrio de oxígeno de la mioglobina cambia poco, pero la hemoglobina se desprende de un 45 por ciento, aproximadamente, del oxígeno que transporta. La diferencia se explica por el efecto amplificador de la interacción hemo-hemo. Pocas moléculas de hemoglobina llevan uno o dos oxígenos. Si la molécula de hemoglobina capta un oxígeno, ésta tiende a continuar y adquirir cuatro oxígenos; si una molécula saturada pierde un oxígeno, se desprenderán, a buen seguro, dos o tres oxígenos más. Los datos para la hemoglobina proceden de las curvas que han sido calculadas por Joyce, M. Baldwin, de Cambridge, Inglaterra.



MAQUINA DE MOVIMIENTO ALTERNANTE, que sirve como modelo de los efectos de cooperatividad de la hemoglobina. El pistón es impulsado a la derecha por la energía liberada en la reacción de la hemoglobina con el oxígeno (O_2) y, hacia la izquierda, por la de los protones (H^+) y dióxido de carbono (CO_2) liberado por los tejidos en la respiración. El difosfoglicerato (DPG) y los iones cloruro (Cl^-) son pasajeros que viajan con los protones y el CO_2.

pista vital nos era desconocida. Como en las novelas de Agatha Christie, la naturaleza lo guardó escondido hasta el final, para hacer la historia más intrigante.

Hay dos maneras de escapar de un callejón sin salida en la ciencia: experimentar o pensar. Por temperamento, quizá, yo experimenté, en tanto que Jacques Monod pensó. Al final, nuestros caminos convergieron.

Monod dedicó su vida científica a investigar qué es lo que regula el crecimiento de las bacterias. La clave de este problema parecía ser la regulación de la síntesis y de la actividad catalítica de los enzimas. Monod y François Jacob habían descubierto que la actividad de ciertos enzimas está controlada por la conexión y desconexión de su síntesis en el gen; ellos y otros encontraron luego un segundo modo de regulación que parecía que operaba activando los enzimas por sí mismos.

En 1965, Monod y Jean-Pierre Changeux, del Instituto Pasteur de París, junto con Jeffries Wyman, de la Universidad de Roma, reconocieron que los enzimas de la última clase, de regulación, tenían ciertas características en común con la hemoglobina. Todos ellos están compuestos por varias subunidades, de manera que cada molécula contiene varios sitios con la misma actividad catalítica, al igual que la hemoglobina contiene varios hemos que enlazan oxígeno y todos ellos muestran efectos de cooperatividad similares. Monod y sus colegas sabían que la desoxihemoglobina y la oxihemoglobina tenían estructuras diferentes, lo que les hizo sospechar que también los enzimas podrían encontrarse en dos (o al menos dos) estructuras. Postularon que tales estructuras debían distinguirse por la disposición de las subunidades y por el número y fuerza de los enlaces entre ellas.

Si hay sólo dos estructuras alternativas, la que mostrara menos, y más débiles, enlaces entre las subunidades podría desarrollar totalmente su actividad catalítica (o afinidad hacia el oxigeno); esta estructura se denominaría *R*, de "relajada". La actividad quedaría amortiguada en la estructura con más, y más fuertes, enlaces entre las subunidades; esta forma se denominó *T*, de "tensa". En cualquiera de estas estructuras, la actividad catalítica (o afinidad hacia el oxígeno) de todas las subunidades en una molécula debe permanecer siempre igual. Este postulado de simetría permitió que las propiedades de los enzimas alostéricos se describieran por medio de una elegante teoría matemática con solamente tres variables independientes: K_R y K_T, las cuales en la hemoglobina representan las constantes de equilibrio del oxígeno de las estructuras *R* y *T* respectivamente, y *L*, que significa el número de moléculas en la estructura *T* dividido por el número en la estructura *R*, midiéndose este cociente en ausencia de oxígeno. Se acuñó el término alostería (de la raíz griega *allos*, "otro", y *stereos*, "sólido") debido a que la molécula reguladora que activa o desactiva al enzima tiene una estructura diferente de la estructura de la molécula cuya transformación química cataliza el enzima.

Esta ingeniosa teoría simplificó enormemente la interpretación de los efectos de cooperatividad. El progresivo incremento en la afinidad para con el oxígeno, ilustrado por la parábola del rico y del pobre, surge ahora, no de una interacción directa entre los hemos, sino de cambios desde la estructura *T* con baja afinidad a la estructura *R* con afinidad elevada. Esta transformación debería tener lugar cuando se enlaza la segunda o la tercera molécula de oxígeno. Los agentes químicos que no se enlazan a los grupos hemo pueden reducir la afinidad para con el oxígeno desplazando el equilibrio entre las dos estructuras hacia la forma *T*, lo que haría que la transición hacia la estructura *R* se produjera, por ejemplo, luego de unir tres moléculas de oxígeno, en vez de dos que eran las que debían enlazarse antes. En el marco de la teoría alostérica, dichos agentes aumentarían *L*, la fracción de moléculas en la estructura *T*, sin alterar las constantes de equilibrio K_T y K_R de las dos estructuras.

Estructuras atómicas

Mi propia investigación sobre el problema también estuvo influida por el descubrimiento de Haurowitz de que la oxihemoglobina y la desoxihemoglobina poseían estructuras diferentes. Gradualmente me fui dando cuenta de que nunca podríamos explicar las intrincadas funciones de la hemoglobina sin resolver las estructuras de ambas formas cristalinas a una resolución lo bastante elevada como para revelar el nivel atómico.

En 1970, 33 años después de que yo hubiera obtenido mi primera imagen de la hemoglobina por difracción de rayos X, se alcanzó finalmente esta fase. Hilary Muirhead, Joyce M. Baldwin, Gwynne Goaman y yo obtuvimos un excelente mapa de la distribución de la materia, no en la oxihemoglobina, sino en la estrechamente relacionada metahemoglobina del caballo; en ella, el hierro se halla en estado férrico y el lugar del oxígeno está ocupado por una molécula de agua. William Bolton y yo conseguimos el mapa de la desoxihemoglobina de caballo, y Muirhead y Jonathan Greer lograron confeccionar otro mapa de desoxihemoglobina humana. Estos mapas sirvieron de guías para la construcción de tres modelos atómicos, cada uno de los cuales semejaba una jungla de varillas de latón y conexiones de acero dispuestas en un bastidor de latón, dando lugar a una construcción de inmensa complejidad de unos 130 centímetros de diámetro. Al principio, los árboles entorpecían la visión del bosque.

En términos alostéricos nuestro modelo de metahemoglobina representaba la estructura *R* y nuestros dos modelos de desoxihemoglobina la estructura *T*. Los analizamos con ahínco en busca de la explicación de los mecanismos alostéricos, pero no dimos con ella al principio porque la estructura general de las subunidades era similar en los tres modelos. Las cadenas alfa incluían siete segmentos helicoidales y las cadenas beta ocho segmentos helicoidales interrumpidos por esquinas y segmentos no helicoidales. Cada cadena envolvía su grupo hemo en una cavidad profunda, expuesta solamente por el borde donde dos cadenas laterales de ácido propiónico de la porfirina se introducían en el agua circundante.

El hemo establece contacto con 16 cadenas laterales de aminoácidos desde siete segmentos de la cadena. La mayoría de estas cadenas laterales son hidrocarbonos; las dos excepciones son las histidinas enlazadas a los grupos hemo, que se disponen a cada lado del hemo y desempeñan un importante papel en la captación del oxígeno. La cadena lateral de la histidina finaliza en un anillo de imidazol compuesto de tres átomos de carbono, dos átomos de nitrógeno y cuatro o cinco átomos de hidrógeno. Una de estas histidinas, denominada la histidina proximal, forma un enlace químico con el hierro del hemo [*véanse las ilustraciones de las páginas 41 y 42*]. La otra histidina, denominada distal, se dispone en el lado opuesto del grupo hemo, en contacto con éste y con el oxígeno enlazado, pero sin formar un enlace químico covalente con ninguno. Fuera de estas histidinas, la mayoría de las cadenas laterales en el interior de las subunidades, tales como aquellas próximas a los hemos, son hidrocarbonos. El exterior de la molécula de la hemoglobina está revestido por cadenas laterales de todas las clases, aunque predominan las cargadas eléctricamente y las dipolares. Así pues, cada subunidad es cérea en el interior y jabonosa en el exterior, lo que las hace solubles en agua aunque impermeables a ésta.

Las cuatro subunidades están dispuestas en los vértices de un tetraedro alrededor de ejes de simetría dobles. Puesto que el tetraedro tiene seis aristas, hay seis áreas de contacto entre las subunidades. La si-

metría doble permite cuatro contactos distintos, que cubren aproximadamente una quinta parte de la superficie de las subunidades. El sesenta por ciento de este área se halla constituido por contactos alfa$_1$-beta$_1$ y alfa$_2$-beta$_2$, cada uno de los cuales abarca unas 35 cadenas laterales de aminoácidos firmemente unidos por enlaces de hidrógeno que oscilan entre 17 y 19. [Los enlaces de hidrógeno están formados entre átomos de nitrógeno (N) y oxígeno (O) por medio de un átomo de hidrógeno (H) intermediario, por ejemplo N–H...N, N–H...O, O–H...O, o, O–H...N. El hidrógeno está enlazado fuertemente al átomo de la izquierda y débilmente al de la derecha.]

Los numerosos enlaces de hidrógeno entre las subunidades alfa$_1$-beta$_1$ y alfa$_2$-beta$_2$ hace que éstas se cohesionen tan fuertemente que sus contactos con dificultad se alteran en su reacción con el oxígeno, y se mueven como cuerpos rígidos en la transición entre las estructuras *T* y *R*. Por otra parte, el contacto alfa$_1$-beta$_2$ en la estructura *R* parece bastante diferente del de la estructura *T*. Este contacto abarca menos cadenas laterales que el alfa$_1$-beta$_1$ y está configurado de manera que actúa como un conmutador de acción rápida, con dos posiciones estables alternativas, cada una ligada por una serie diferente de enlaces de hidrógeno. Nos preguntamos al principio si estos enlaces serían más fuertes y numerosos en la estructura *T* que en la *R*, pero éste no parecía ser el caso.

¿Dónde estaban, pues, los enlaces extra entre las subunidades en la estructura *T* que exigía la teoría alostérica? Los supusimos en los extremos finales de las cadenas polipeptídicas. En la estructura *T* el último resto de aminoácido de cada cadena forma puentes salinos con las subunidades vecinas. (Un puente salino es un enlace entre un átomo de nitrógeno, que transporta una carga positiva, y un átomo de oxígeno, que lleva una carga negativa.) En nuestros mapas de la estructura *R* los dos últimos restos de cada cadena eran confusos. Al principio sospeché que ello se debía a un error, pero mapas perfeccionados realizados por mis colegas Elizabeth Heidner y Robert Ladner nos convencieron de que los restos finales permanecían invisibles porque no están bien sujetos y ondean de forma parecida a como lo hacen las espigas con el viento.

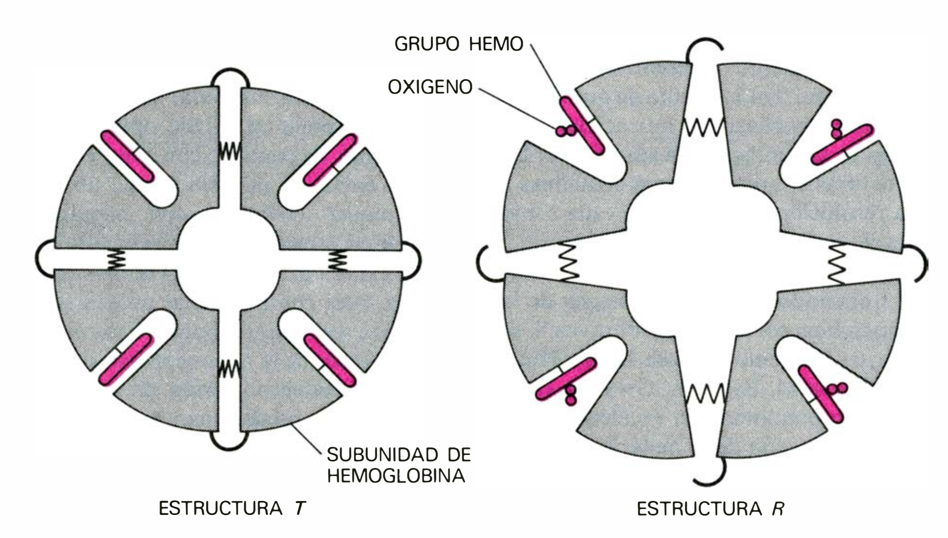
Geométricamente, la transición entre las dos estructuras consiste en un movimiento de "rock-and-roll" del dímero alfa$_1$-beta$_1$ con respecto al dímero alfa$_2$-beta$_2$. Baldwin ha demostrado que si un dímero se mantiene fijo, el movimiento del otro puede ser representado por una rotación de unos 15 grados en un eje colocado convenientemente junto con un pequeño desplazamiento a lo largo del mismo eje. El movimiento viene inducido por sutiles cambios en la estructura interna de las subunidades que acompañan al enlace y disociación del oxígeno.

Función de los puentes salinos

Los puentes salinos en los extremos de las cadenas polipeptídicas proporcionan claramente los enlaces adicionales entre las subunidades en la estructura *T* predicha por Monod, Changeux y Wyman. Explican también la influencia en la curva de equilibrio del oxígeno de todos los factores químicos que nos han desconcertado tanto. Todos los agentes que rebajan la afinidad para con el oxígeno lo hacen reforzando los puentes salinos existentes en la estructura *T*, o añadiendo otros nuevos. Pero no todos estos enlaces adicionales se hallan entre las subunidades; algunos están en el interior de las subunidades y se oponen a los sutiles cambios estructurales que las subunidades llevan a cabo en combinación con el oxígeno.

Los puentes salinos explican tanto el descenso en la afinidad de oxígeno por los protones como la captación de protones en la liberación de oxígeno. Los protones aumentan el número de átomos de nitrógeno que transportan una carga positiva. Por ejemplo, el anillo imidazol del aminoácido histidina puede presentarse en dos estados: sin carga, cuando sólo uno de sus átomos de nitrógeno transporta un protón, y cargado positivamente cuando ambos lo hacen. En solución neutra, cada histidina tiene una posibilidad del 50 por ciento de estar cargada positivamente. Cuanto más ácida es la solución, o, en otras palabras, cuanto más alta es la concentración de protones, mayor es la posibilidad de que la histidina se cargue positivamente y forme un puente salino con un átomo de oxígeno que transporte una carga negativa. A la inversa, la transición de la estructura *R* a la estructura *T* pone en contacto átomos de oxígeno cargados negativamente con átomos de nitrógeno sin carga, y, por tanto, disminuye el trabajo requerido para que el átomo de nitrógeno obtenga una carga positiva. El resultado es que la histidina, que no tiene una posibilidad mayor del 50 por ciento de estar cargada positivamente en la estructura *R*, alcanza una posibilidad del 90 por ciento en la *T*, de manera que la hemoglobina capta más protones de la solución en la estructura *T*.

La hemoglobina comprende otra serie de grupos que se comportan de este modo: son los grupos amino del principio de las cadenas polipeptídicas, aunque sus átomos de nitrógeno sólo captan protones si la concentración de dióxido de carbono es baja. Si es elevada, estos nitrógenos son propensos a perder protones y a combinarse con el dióxido de carbono para formar un compuesto carbamínico. Los fisiólogos F. J. W. Roughton y J. K. W. Ferguson propusieron, en 1934, que este mecanismo intervenía en el transporte del dióxido de carbono, pero sus ideas se recibieron con escepticismo hasta que las confirmó 35 años más tarde mi colega John Kilmartin, que trabajaba con Luigi Rossi-Bernardi en la Universidad de Milán.

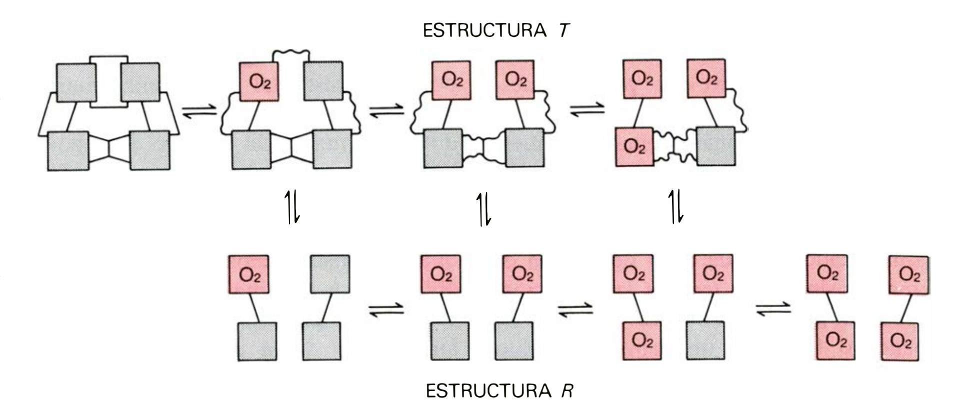


LA TEORIA ALOSTERICA explica la interacción hemo-hemo sin postular ninguna comunicación directa entre los grupos hemo. Se supone que la molécula de la hemoglobina tiene dos estructuras alternativas, designadas *T* (tensa) y *R* (relajada). En la estructura *T*, las subunidades de la molécula están sujetas en contra de la presión de unos resortes y sus estrechas cavidades impiden la entrada del oxígeno. En la estructura *R*, todas las sujeciones han cedido abriéndose, y las cavidades de los grupos hemo se abren para admitir sin dificultad al oxígeno. La captación de oxígeno por la estructura *T* estiraría las sujeciones hasta que se abrieran, repentinamente, todas juntas permitiendo que la molécula se relajara hasta la estructura *R*. La pérdida de oxígeno estrecharía las cavidades de los hemo y permitiría que se formara de nuevo la *T*.

Me alegré de que Roughton, que fue el precursor del experimento, viviera aún para ver su idea reivindicada. Mi colega Arthur R. Arnone, quien se halla ahora en la Universidad de Iowa, demostró entonces que en la estructura *T* tales grupos carbamínicos, que transportan cargas negativas, forman puentes salinos con grupos cargados positivamente de la globina, y son, por consiguiente, más estables que en la estructura *R*. Este hallazgo explica la razón de que la desoxihemoglobina muestre una afinidad más elevada hacia el dióxido de carbono que la oxihemoglobina, y, a la inversa, justifica por qué el dióxido de carbono disminuye la afinidad para con el oxígeno de la hemoglobina.

Siguen sin conocerse los lugares en que los iones cloruro se engarzan en la hemoglobina. Arnone ha marcado lugares en la estructura *T* donde se enlazan otros iones cargados negativamente, y quizá sean éstos también los lugares de enlace del cloruro. Si realmente lo son, los iones cloruro deberán unirse a la estructura *T* formando puentes salinos adicionales.

La diferencia más espectacular entre las estructuras *T* y *R* es la amplitud del espacio entre las dos cadenas beta. En la estructura *T*, las dos cadenas se hallan muy separadas, y el vacío que media entre ellas está recubierto por cadenas laterales de aminoácidos que llevan cargas positivas. Este vacío está "hecho a medida" para que albergue a la molécula de 2,3-difosfoglicerato y compense sus cargas negativas; así pues, el enlace con el DPG añade otra serie de puentes salinos a la estructura *T*. En la estructura *R*, la abertura se estrecha, y el DPG tiene que desprenderse.



TRANSICION de la estructura *T* a la estructura *R*. Tiene por consecuencia incrementar sus posibilidades a medida que cada uno de los cuatro grupos hemo se va oxigenando. En este modelo más realista, los puentes salinos que unen las subunidades en la estructura *T* se rompen progresivamente a medida que se incorpora el oxígeno, y hasta los puentes salinos que no se han roto aún se debilitan, proceso que aquí esta representado por un trazo de línea ondulada. La transición de *T* a *R* no tiene lugar hasta que un número fijo de moléculas de oxígeno se ha unido, pero ésta se hace más probable con cada oxígeno que se va enlazando. En la transición entre ambas estructuras intervienen varios factores: protones, dióxido de carbono, cloruro y DPG. Cuanto mayor es su concentración, más oxígeno debe unirse para desencadenar la transición. Las moléculas totalmente saturadas en la estructura *T* y las moléculas totalmente desoxigenadas en la estructura *R* no se muestran porque resultan demasiado inestables para encontrarse en números significativos.

El desencadenante

¿De qué manera la combinación del hierro del hemo con el oxígeno hace que las subunidades cambien inmediatamente de la estructura *T* a la estructura *R*? Comparada con la molécula de hemoglobina, la molécula de oxígeno es como una pulga que obliga a bailar a un elefante. A la inversa, ¿cómo se las ingenia la estructura *T* para impedir la captación de oxígeno? ¿Qué diferencia entre ambas estructuras se registra en el hemo, la cual puede conducir a un cambio en la afinidad para con el oxígeno de varios cientos de veces mayor o menor?

En la oxihemoglobina, el hierro del grupo hemo está enlazado a seis átomos: a cuatro átomos de nitrógeno de la porfirina, que neutralizan las dos cargas positivas del hierro en estado ferroso; a un átomo de nitrógeno de la histidina proximal, que enlaza el hemo con uno de los segmentos helicoidales de la cadena polipeptídica (hélice *F*); y a uno de los dos átomos de la molécula de oxígeno. En la desoxihemoglobina la posición del oxígeno permanece vacía, de modo que el hierro se enlaza sólo a 5 átomos.

Me preguntaba si las cavidades de los hemos podrían ser más estrechas en la estructura *T* que en la estructura *R*, de suerte que tuvieran que ensancharse para dar cabida al oxígeno. Este ensanchamiento podría estar conectado con la ruptura de los puentes salinos, de forma parecida al ingenuo mecanismo mostrado en la figura de la página opuesta. Cuando se consiguió el modelo atómico de la desoxihemoglobina de caballo, Bolton y yo comprobamos parte de esta idea, puesto que, en las subunidades beta, una cadena lateral del aminoácido valina inmediato a la histidina distal bloqueaba el lugar que hubiera ocupado el oxígeno. Las subunidades alfa, sin embargo, no mostraron tal obstrucción. Luego nos dimos cuenta de las extrañas posiciones de los átomos de hierro. En la metahemoglobina, que tiene una estructura en *R*, los átomos de hierro han sido desplazados muy ligeramente del plano de la porfirina hacia la histidina proximal, si bien en la desoxihemoglobina (con la estructura en *T*) el desplazamiento se mantenía como una de las características más notables de nuestros mapas. En cada subunidad, el átomo de hierro había transportado la histidina proximal y la hélice *F* con él, de modo que también éstos se alejaron del plano de la porfirina. En seguida me di cuenta de que éste podía ser el desencadenador buscado desde hacía tanto tiempo.

Recientemente, Arnone y mi colega Lynn Ten Eyck han obtenido un excelente mapa de la desoxihemoglobina humana al que Giulio Fermi ha ajustado un modelo atómico del grupo hemo empleando métodos cibernéticos. Los cálculos de Fermi muestran que cada átomo de hierro está desplazado 0,6 (± 0,1) angstrom unidad del plano principal de la porfirina. (Un angstrom unidad equivale a 10^{-10} metros.) El átomo de nitrógeno de la histidina proximal, que está enlazado al átomo de hierro, se dispone a una distancia de 2,7 (± 0,1) angstrom del mismo plano. No poseemos todavía una medida directa de los desplazamientos correspondientes en la oxihemoglobina, porque ésta se oxida a metahemoglobina mediante la acción de los rayos X. Allí los átomos de hierro se desplazan del plano de la porfirina 0,1 angstrom en las subunidades alfa y 0,2 angstrom en las subunidades beta; los correspondientes desplazamientos de los nitrógenos de la histidina son 2,2 y 2,4 angstrom. A juzgar por las estructuras de los modelos compuestos, el desplazamiento del nitrógeno de la histidina en la oxihemoglobina debe ser de 2,1 angstrom, lo que significa que el nitrógeno tendría que estar 0,6 angstrom más cerca del plano de la porfirina que como se halla en la desoxihemoglobina. Este desplazamiento podría desencadenar la transición de la estructura *T* a la estructura *R*.

¿Cómo se transmite este movimiento a los contactos entre las subunidades y a los puentes salinos? Al igual que podemos deducir cómo salta un gato la pared a partir de fotografías en las que el animal primero está sobre el muro y luego

en las que ya se encuentra en el suelo, hemos de intentarlo aquí porque nuestros modelos estáticos de desoxihemoglobina y metahemoglobina no muestran qué es lo que pasa en la transición entre las estructuras *T* y *R*. Formulé una conjetura atrevida. El segundo resto de aminoácido de cada final de cadena es una tirosina cuya cadena lateral transporta un grupo fenol, o sea, un anillo de benceno ligado a un grupo hidroxilo (OH). En la estructura *T*, la tirosina de cada subunidad está acuñada en una cavidad entre las hélices *F* y *H* y su grupo hidroxilo está trabado por un enlace de hidrógeno a un átomo de oxígeno en el segmento FG de la cadena polipeptídica principal. En la carboxihemoglobina, que es el pariente más cercano de la oxihemoglobina y que tiene una estructura *R*, las tirosinas están libres. En consecuencia, debe haber un mecanismo que desprenda las tirosinas cuando el oxígeno entre en combinación.

Andaba yo preguntándome en qué podría consistir tal mecanismo, cuando observé que el movimiento de la histidina proximal hacia el plano de la porfirina que acompaña al enlace con el oxígeno empujaba la hélice *F* en una dirección que estrechaba la cavidad en la que la tirosina debía instalarse. Si se obligara a la tirosina a salir fuera de su cavidad, desgarraría el último resto de aminoácido de la cadena separándolo del otro elemento del puente salino. Se podría romper así un puente salino por cada grupo hemo que se combinara con oxígeno en la estructura *T*. Una vez rotos suficientes puentes salinos, la estructura *T* se volvería inestable y rápidamente cambiaría a la estructura *R*.

Si el movimiento de la histidina proximal y del hierro hacia la porfirina pone en movimiento una serie de palancas que desprenden las tirosinas y rompen los puentes salinos, entonces la composición de los puentes y el engarzamiento de las tirosinas en sus cavidades deben causar que la misma serie de palancas funcionen en sentido inverso y desplacen la histidina y el hierro lejos de la porfirina. La molécula de oxígeno, por otra parte, no puede seguirlos, porque se encuentra aprisionada por los cuatro átomos de nitrógeno porfirínico, de manera que el enlace del hierro con el oxígeno se mantiene hasta el final, en que se rompe.

Seguir el camino, de la mano de los modelos atómicos, que llevaba hacia el mecanismo molecular del transporte respiratorio parecía todo un sueño. O, ¿acaso era algo real? ¿Aguantaría el mecanismo el frío análisis de un experimento? Se ha dicho que los científicos no persiguen la verdad, ésta les persigue a ellos.

EJE DE SIMETRIA
ALFA$_2$
ALFA$_1$
EJE DE ROTACION
BETA$_1$
BETA$_2$
15 GRADOS
7,5 GRADOS

REESTRUCTURACION DE LAS SUBUNIDADES durante la transición de la estructura *T* a la estructura *R*. Dicha reordenación consiste, fundamentalmente, en la rotación de un par de subunidades con respecto al otro par. Cada cadena alfa está fuertemente unida a una cadena beta, y el dímero así formado se mueve como un cuerpo rígido. Si un dímero se mantiene fijo, el otro gira 15 grados alrededor de un eje descentrado y se desplaza un tanto a lo largo de éste. La simetría doble de la molécula se mantiene, si bien el eje de simetría ha girado 7,5 grados. El diagrama se funda en otro anterior realizado por Joyce M. Baldwin.

Sometiendo a prueba los puentes salinos

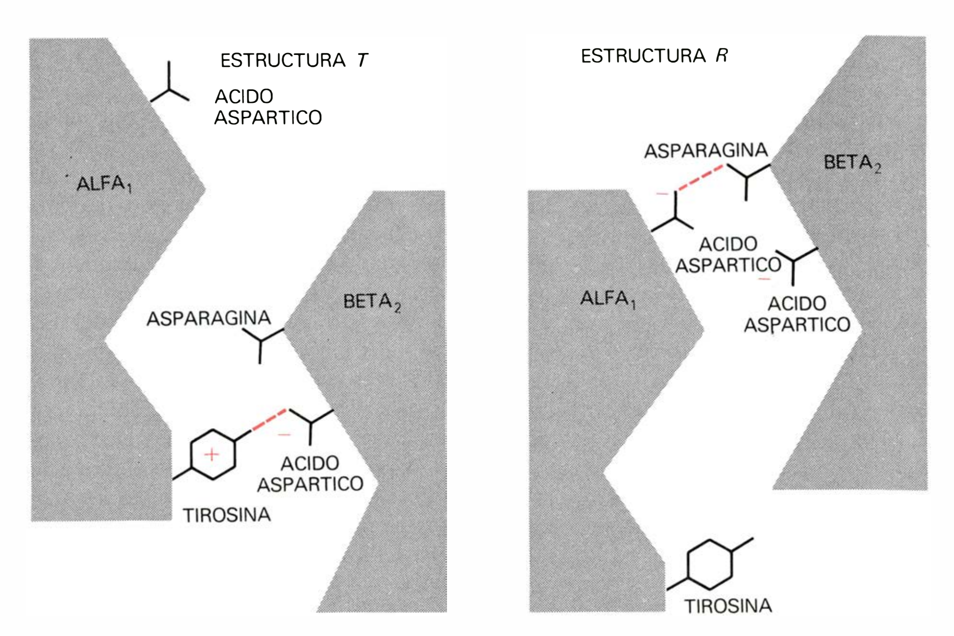
Según la teoría alostérica, no debe haber interacción hemo-hemo sin transición entre las estructuras *T* y *R*. Esta predicción también fue comprobada por Kilmartin. Eliminó el aminoácido final de los extremos de las cuatro cadenas polipeptídicas, de manera que no podía haber puentes salinos que estabilizaran la estructura *T*. Esta hemoglobina modificada mantuvo la estructura *R*, incluso en ausencia de oxígeno, y mostró una curva de equilibrio de oxígeno hiperbólica con una elevada afinidad para con el oxígeno. Kilmartin seleccionó luego una hemoglobina humana anormal que se podía hacer que mantuviera la estructura *T* incluso cuando estaba saturada con oxígeno. De nuevo resultó una curva hiperbólica, si bien estaba desplazada hacia una afinidad de oxígeno menor, de forma que pudo ser demostrada la tesis central de la alostería.

La cuestión siguiente se refería al papel exacto de los puentes salinos. En un extremo, la tensión originada por la combinación de uno cualquiera de los grupos hemo con el oxígeno podría distribuirse de manera uniforme por toda la molécula; no habría, pues, cambio en la afinidad de oxígeno hasta que todos los puentes salinos se rompieran al unísono, cuando la estructura cambiara de *T* a *R*. Esto encajaría con la teoría alostérica pura, según la cual lo único que deben hacer los puentes salinos es aumentar *L*, la fracción de moléculas en la estructura *T*. Por otra parte, según mi mecanismo, tiene que romperse un puente salino cada vez que un oxígeno se combina con la estructura *T*. Si eso fuera cierto, los puentes salinos deberían

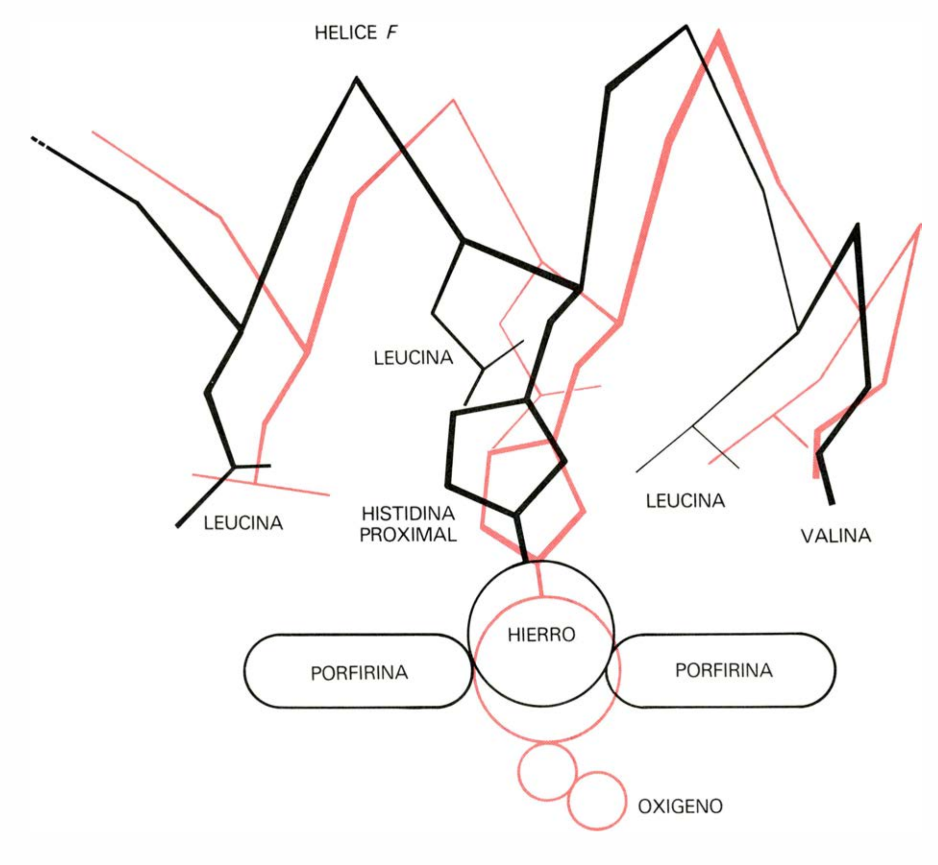
aumentar K_T además de L, es decir, rebajarían asimismo la afinidad de la estructura T para con el oxígeno. Al lector no especializado esto puede parecerle una distinción de matiz; pero si fuera un especialista en la materia quizá se rasgase las vestiduras ante lo que parecería cortar de raíz los mecanismos establecidos; y, efectivamente, se levantaron enconadas controversias que todavía perduran.

Varios resultados experimentales apoyaron mi versión del mecanismo. En 1965, Eraldo Antonini, Todd M. Schuster, Maurizio Brunori y Jeffries Wyman, de la Universidad de Roma, y, en 1970, R. D. Gray, de la Universidad Cornell, demostraron que la captación de oxígeno y la liberación de los protones Bohr van juntas desde el principio, mientras la hemoglobina sigue en la estructura T. Kilmartin demostró luego que la mayoría de los protones Bohr provienen de la ruptura de los puentes salinos. Juntos, ambos resultados prueban que los puentes salinos se rompen al captar oxígeno la estructura T, lo que implica que éstos rebajan la afinidad del oxígeno. Para comprobar si tal cosa ocurre realmente tuvimos que hacer una comparación precisa de las curvas de equilibrio del oxígeno de la hemoglobina normal y de la hemoglobina que carece de uno de los puentes salinos. Kiyohiro Imai e Hideki Morimoto, de la Universidad de Osaka, acababan de desarrollar un ingenioso método que permitía que la curva de equilibrio del oxígeno se midiera con gran precisión y rapidez con sólo 0,1 mililitros de solución de hemoglobina. Imai vino a Cambridge para montar uno de sus nuevos aparatos y, con Kilmartin, midió las curvas de equilibrio de las hemoglobinas que carecían de puentes salinos específicos. Observaron que la ausencia de cualquiera de los puentes dejaba K_R invariable, pero disminuía tanto L como K_T, en perfecto acuerdo con mi mecanismo.

Paradójicamente, otra serie de observaciones contradice estos hallazgos. Mis colegas Leigh Anderson y Kilmartin, junto con Seiji Ogawa, de los Laboratorios Bell, han mostrado que los puentes salinos se rompen únicamente si la molécula de hemoglobina está libre para pasar a la estructura R, pero no si esta transición se frena. Esto ocurre en cierta hemoglobinas humanas anormales y en las hemoglobinas de peces en disolución ácida, donde la estructura T es extremadamente estable. Parece que la estructura T debe estar libre para doblarse y extenderse como si se tuviera que agitar hasta liberarse de sus impedimentos;



CONTACTO entre los dos dímeros. Tiene dos formas estables, una para la estructura *T* y otra para la estructura *R*. En la transición entre ambas estructuras, los dímeros pasan repentinamente de una posición a otra. Las formas se hallan estabilizadas por series alternativas de enlaces de hidrógeno entre las cadenas laterales de aminoácidos engarzados en las superficies opuestas de los dímeros. Los dos enlaces que se muestran aquí fueron descubiertos, por primera vez, mediante cristalografía de rayos X. En 1975, Leslie Fung y Chien Ho demostraron la presencia de estos enlaces en solución. Esto nos aporta la prueba de que las dos estructuras halladas en cristales corresponden a las mismas estructuras de los glóbulos rojos.

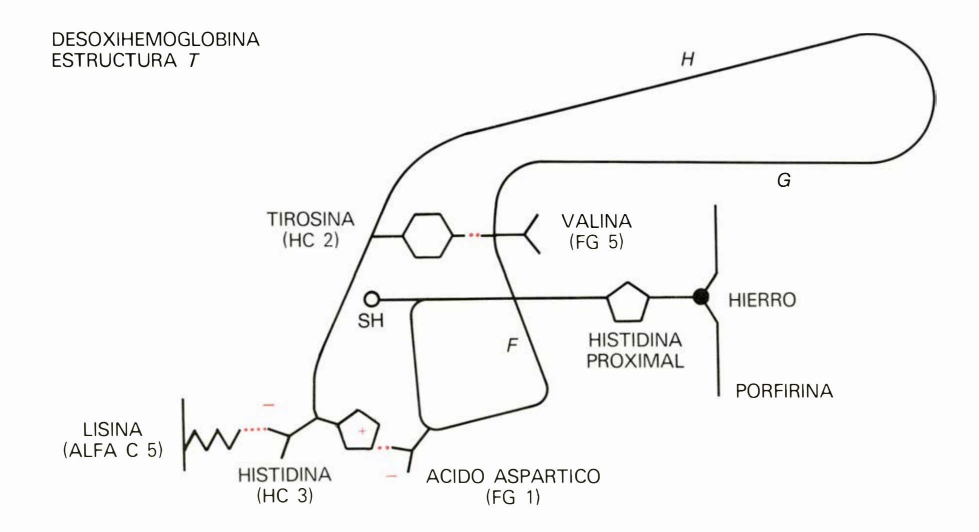


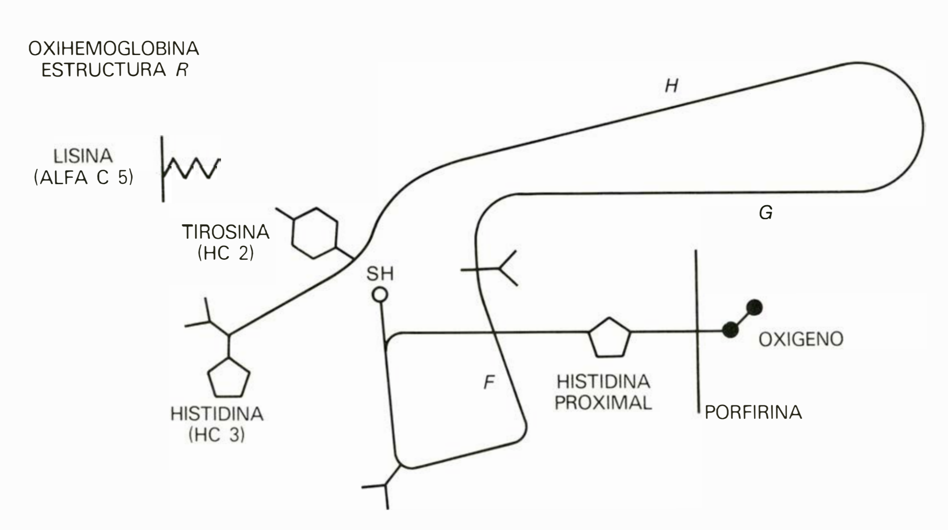
MECANISMO DE DESENCADENAMIENTO de la transición de la estructura *T* a la *R*. Se trata de un movimiento del hierro hemo en el plano del anillo de la porfirina. En la estructura *T (líneas negras)*, el centro del átomo de hierro está a unos 0,6 angstrom por encima del plano (un angstrom equivale a 10^{-10} metros). Cuando la molécula pasa a la estructura *R (líneas en color)*, el hierro se mueve dentro del plano, estirando con él la histidina proximal y la hélice *F*. Una vez que el hierro ha descendido en el plano puede enlazarse rápidamente con una molécula de oxígeno. En la transición inversa (de *R* a *T*) se desplaza el hierro fuera del plano; el oxígeno no puede seguirle porque tropieza con los átomos de nitrógeno porfirínico. El enlace hierro-oxígeno se debilita, por consiguiente, y acaba por romperse. Estos movimientos se transmiten a los contactos entre las subunidades y facilitan las transiciones entre las estructuras *T* y *R*. El diagrama de la ilustración está basado en un dibujo realizado por John Cresswell, del University College de Londres.

cuando permanece ligada con excesiva intensidad deja de responder.

He sugerido que la transición de la estructura *T* a la *R* viene desencadenada principalmente por el movimiento del hierro del grupo hemo hacia el anillo de la porfirina. ¿Qué determina que el hierro se mueva? Hay dos razones, una espacial y la otra electrónica. Si el hierro está enlazado a átomos a ambos lados del hemo, la atracción de los átomos hacia el hierro y la repulsión de éste por los nitrógenos porfirínicos tenderán a equilibrar el hierro en el centro del anillo. Por otra parte, si el hierro está enlazado sólo a la histidina proximal y el lugar del oxígeno se halla vacío, entonces la repulsión entre los nitrógenos porfirínicos y la histidina no se equilibra por la repulsión entre los nitrógenos porfirínicos y el oxígeno, de forma que la histidina se aleja de la porfirina y arrastra al hierro con ella.

La cuestión electrónica es más compleja. El átomo de hierro en estado ferroso tiene seis electrones en su capa más externa. En la oxihemoglobina éstos forman tres pares localizados a medio camino entre los enlaces que unen el hierro con los seis átomos circundantes. La repulsión entre los electrones del hierro y los electrones de los átomos circundantes está, por consiguiente, minimizada. En la desoxihemoglobina, por otra parte, cuatro de los seis electrones no están apareados y dos se sitúan a lo largo del enlace, donde repelen los átomos circundantes del anillo de la porfirina. Esta repulsión tiende a empujar al hierro más lejos del plano de la porfirina que lo que se hubiera producido por la repulsión entre la histidina proximal y los átomos de nitrógeno porfirínico.





LOS ENLACES ADICIONALES en la desoxihemoglobina están formados por los dos últimos restos de la cadena beta. En la oxihemoglobina, el átomo de hierro se dispone en el plano de la porfirina, el grupo sulfhidrilo (SH) del aminoácido cisteína se sitúa en la cavidad entre las hélices *F* y *H*, y la tirosina *HC* 2, la histidina *HC* 3 y la lisina *C* 5 están libres. En la desoxihemoglobina, el hierro se encuentra desplazado del plano de la porfirina hacia la hélice *F*, y la tirosina ha desplazado al grupo SH de la cavidad entre *F* y *H* y forma un enlace de hidrógeno con el aminoácido valina *FG* 5. Por último, la histidina terminal forma puentes salinos con el ácido aspártico *FG* 1 de la misma cadena y con la lisina *C* 5 de la cadena alfa. La formación del puente hace que el aminoácido histidina capte un protón y pase a estar cargada positivamente.

Comprobando el desencadenante

Supóngase que el hierro se mueve hacia dentro y hacia fuera del plano de la porfirina cada vez que se combina o que se pierde una molécula de oxígeno. ¿Cómo podríamos averiguar si es éste el movimiento real que desencadena la transición alostérica entre las dos estructuras? No se me ocurría ningún experimento que contestara a la pregunta directamente, pero me dije que si mi proposición era cierta, entonces, en virtud de las leyes de acción y reacción, una transición forzada desde *R* a *T* debería hacer marcha atrás y arrastrar el hierro y la histidina fuera del anillo porfirínico. En este caso, la estructura *T* debería ejercer una tensión en el grupo hemo, que tendría que ser detectable por métodos físicos. Mi profesor David Keilin me recomendó siempre que trabajara con proteínas coloreadas, puesto que el espectro de la luz que éstas absorben pueden revelar muchas cosas. La hemoglobina resulta doblemente bendecida porque puede apreciarse su pulso tanto por su espectro de absorción como por las propiedades magnéticas de sus átomos de hierro.

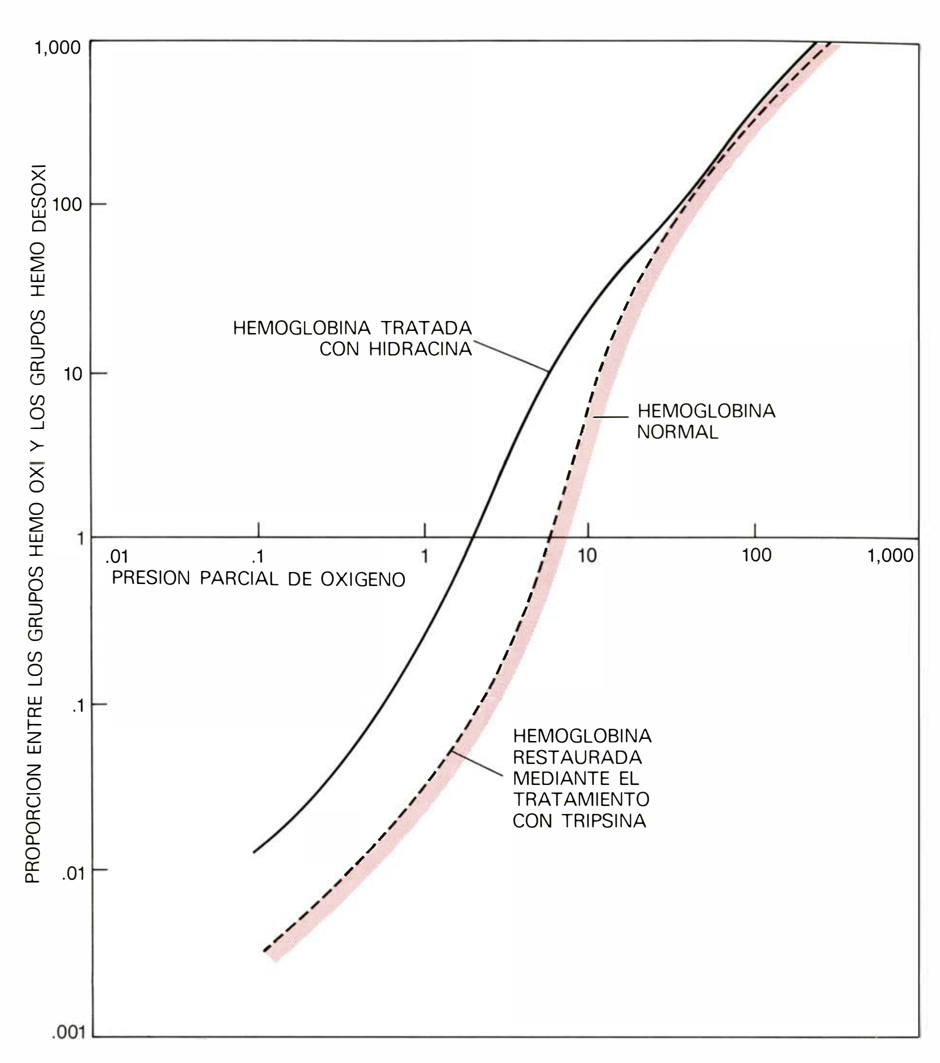
Antes de que pudiéramos explotar estas propiedades teníamos que hallar un método para pasar de la estructura *R* a la *T*, diferente del usual de extraer el oxígeno. Sanford R. Simon, de la Universidad del estado de Nueva York en Stony Brook, y yo encontramos que este paso podía verificarse con una sustancia análoga al DPG, denominada inositol hexafosfato (IHP), que tiene seis grupos fosfato en lugar de los dos del DPG, y, por consiguiente, se enlaza con la estructura *T* de un modo más fuerte.

Cuando se añadió IHP a la hemoglobina, aquél ocasionó que parte del oxígeno se desprendiera, como era de esperar. Reemplacé luego el oxígeno por óxido nítrico (NO), ya que este gas se combina con el hierro tan fuertemente que el enlace, una vez formado, no puede romperse. Cuando añadí IHP a la hemoglobina con óxido nítrico, la estructura viró de *R* a *T*, y el espectro cambió de una manera brusca. Los análisis de estos y de otros cambios del espectro nos revelaron lo que había ocurrido: debido a que el fuerte enlace del óxido nítrico había mantenido al átomo de hierro apretado al plano de la porfirina, la tensión ejercida por la estructura *T* había roto el débil enlace entre el hierro y la histidina proximal. Lo más sor-

prendente es que esto se había producido fundamentalmente en las subunidades alfa, cuyos grupos hemo distan 35 angstrom del lugar de enlace de los fosfatos, en vez de acontecer en las subunidades beta, a las que realmente está unido el IHP. Este experimento lo llevamos a cabo Kyoshi Nagai, Attila Szabo y yo en Cambridge, junto con John C. Maxwell y Winslow S. Caughey, de la Universidad del estado de Colorado en Fort Collins. Robert Cassoly, del Instituto de Biología Físicoquímica de París, descubrió los cambios espectrales al mismo tiempo que nosotros.

Nuestro experimento demostró que la tensión existe aunque no nos reveló su magnitud. Para medir ésta decidí aprovechar ciertos compuestos de hemoglobina en los que los átomos de hierro se encuentran en estado de equilibrio entre un estado paramagnético débil y otro fuerte. (Una sustancia paramagnética no puede estar permanentemente magnetizada, como lo puede estar el hierro metálico, si bien es atraída hacia un campo magnético.) A temperatura baja todos los átomos de hierro son débilmente paramagnéticos: El paramagnetismo desciende a medida que la temperatura aumenta; por encima de cierta temperatura, los átomos de hierro empiezan a oscilar entre los dos estados magnéticos, lo que produce que el paramagnetismo total aumente a medida que sube la temperatura. Se sabe hoy que los enlaces entre el hierro y sus átomos circundantes son un tanto más largos en el estado paramagnético fuerte que en el estado paramagnético débil. Por consiguiente, si la tensión en la estructura *T* alarga los enlaces del hierro, ésta debería determinar una concentración de átomos de hierro en el estado fuertemente paramagnético mayor en la estructura *T* que en la estructura *R*, y, por tanto, aumentar el paramagnetismo total de la solución.

Tras fallidos intentos iniciales, una feliz coincidencia encarriló por fin este experimento. Robert W. Noble se presentó en Cambridge, procedente de la Universidad del estado de Nueva York en Buffalo, con sus bolsillos llenos de hemoglobina de carpa. Me enseñó con cuánta facilidad la estructura de cualquiera de los derivados de esta hemoglobina podían pasar desde *R* a *T* añadiendo una pequeña cantidad de ácido e IHP. Juntos nos pusimos en camino hacia Roma, donde Massimo Cerdonio y Calogero Messana acababan de montar un magnetómetro superconductor extremadamente sensible en el Laboratorio Snamprogetti, pero mientras hacíamos trasbordo en el metropolitano londinense,

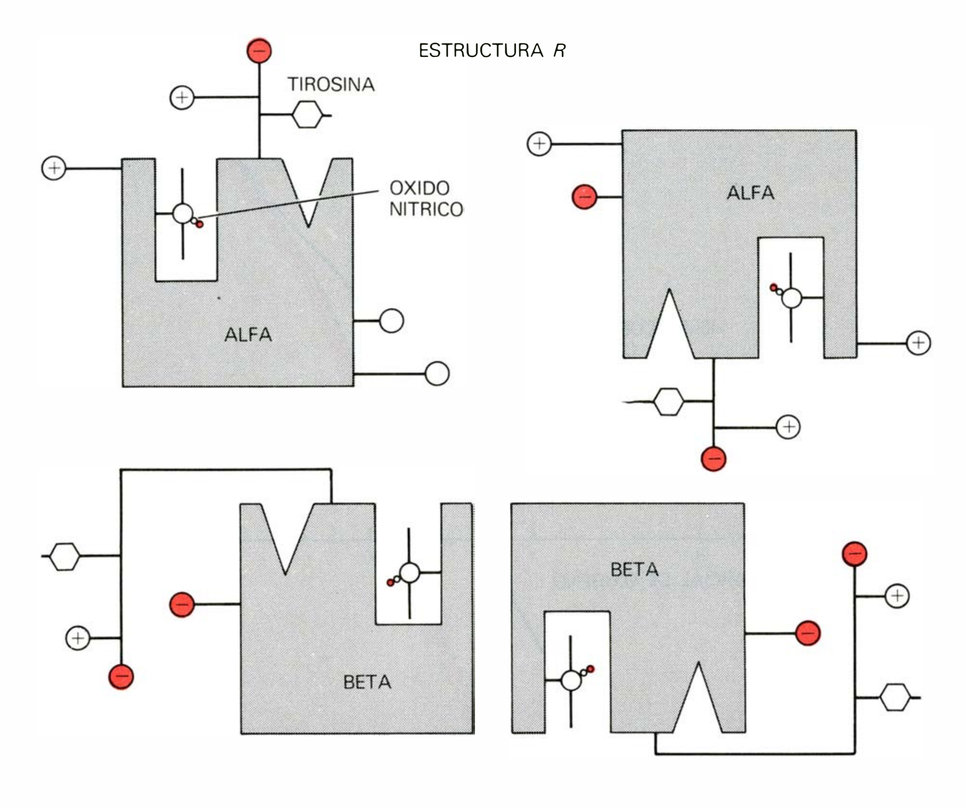


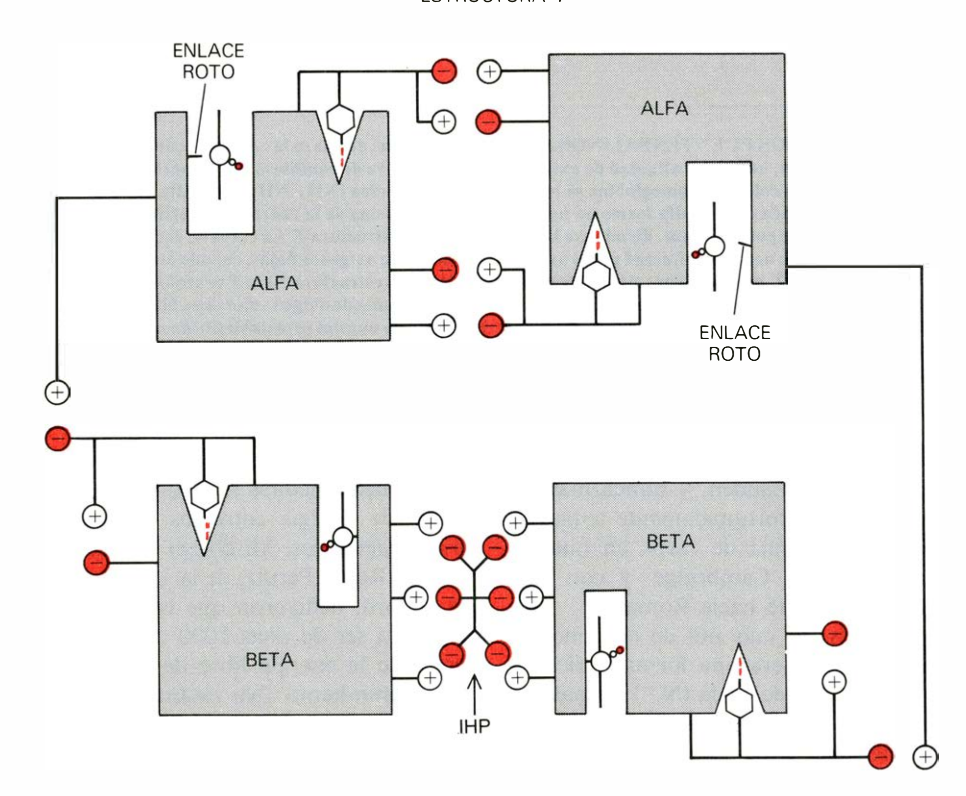
BLOQUEO DE LOS PUENTES SALINOS en la posición abierta; desplaza la curva de equilibrio del oxígeno a la izquierda, hacia una afinidad de oxígeno mayor. La curva de equilibrio inicial para la hemoglobina se muestra en color. La hemoglobina se trató luego con hidracina (NH_2-NH_2), que evita que el carboxilo terminal de cada cadena alfa forme un puente salino con la lisina de la cadena alfa vecina. Cuando el puente salino no se puede formar, disminuye la estabilidad de la estructura *T*. La curva de equilibrio resultante se muestra en negro: la afinidad por el oxígeno a presiones de oxígenos bajas, cuando la hemoglobina tiene la estructura *T*, aumenta unas cinco veces y la transición de la estructura *T* a la *R* se produce antes en la reacción. (Tanto la K_T como la *L* se reducen.) No hay efecto a presiones de oxígeno elevadas, donde las moléculas normales y las tratadas están en la estructura *R*. El equilibrio normal se restableció *(línea discontinua)* tratando la proteína alterada con tripsina, la cual extrae la hidracina. El experimento fue llevado a cabo por John Kilmartin y Janice Fogg, de Cambridge, y Arthur R. Arnone, de la Universidad de Iowa (USA).

dejé los termos con nuestras preciosas muestras en el andén, y nunca más las volví a ver. Afortunadamente teníamos más hemoglobina de carpa en nuestro congelador de Cambridge, y con ésta salí nuevamente hacia Roma.

El derivado más útil de la hemoglobina de carpa era una forma férrica en la que un ion de azida (N_3^-) ocupaba el lugar del oxígeno. Medimos su paramagnetismo tanto en la estructura *R* como en la estructura *T* entre −180 y + 30 grados Celsius. Los resultados nos produjeron una tremenda emoción: a cualquier temperatura, la metahemoglobina azida de la carpa poseía un paramagnetismo mucho más fuerte en la estructura *T* que en la estructura *R*; ello demostraba que la estructura *T* favorecía el estado del hemo con enlaces más largos hierro-nitrógeno. La tensión en el hemo puede medirse a partir de la diferencia de energía entre los dos equilibrios magnéticos. Mi colega Fermi, con mi hijo Robin Perutz, de la Universidad de Oxford, dedujeron que la diferencia llegaba a ser de unas 1000 calorías, un tercio de la energía libre de la interacción hemo-hemo. No estamos seguros de dónde provienen los dos tercios restantes, aunque sospechamos que la transición de *R* a *T* produce un cambio más pequeño en la estructura del hemo, y, por consiguiente, un cambio también menor de tensión en la metahemoglobina azida del que ocurre en la captación y pérdida de oxígeno.

Mientras tanto, Arieh Warshel, de la Universidad Southern de California, Bruce W. Gelin y Martin Karplus, de la





LA TENSION en la estructura *T* se demostró mediante una transición forzada de la estructura *R* a la *T*. Se descubrió ese efecto cuando el autor experimentaba el efecto de la transición en todos los derivados de la hemoglobina. La transición se forzó con la adición de inositol hexafosfato (IHP), que reemplaza al DPG aunque forma más puentes salinos con las cadenas beta. El oxígeno se reemplazó por óxido nítrico, el cual se enlaza al hierro muy fuertemente y debilita, a su vez, el enlace hierro-histidina. En la parte superior, todos los hierros hemos se han unido a NO, las moléculas están en la estructura *R* y no hay puentes salinos entre las subunidades. En la parte inferior, el IHP ha convertido a la molécula a la estructura *T* y las subunidades están sujetas por los puentes salinos. La tensión resultante rompe los enlaces entre los átomos de hierro y las histidinas proximales en las subunidades alfa, que están mucho más lejos de los lugares de enlace del IHP que los de las beta: las proteínas pueden-transmitir efectos mecánicos a grandes distancias.

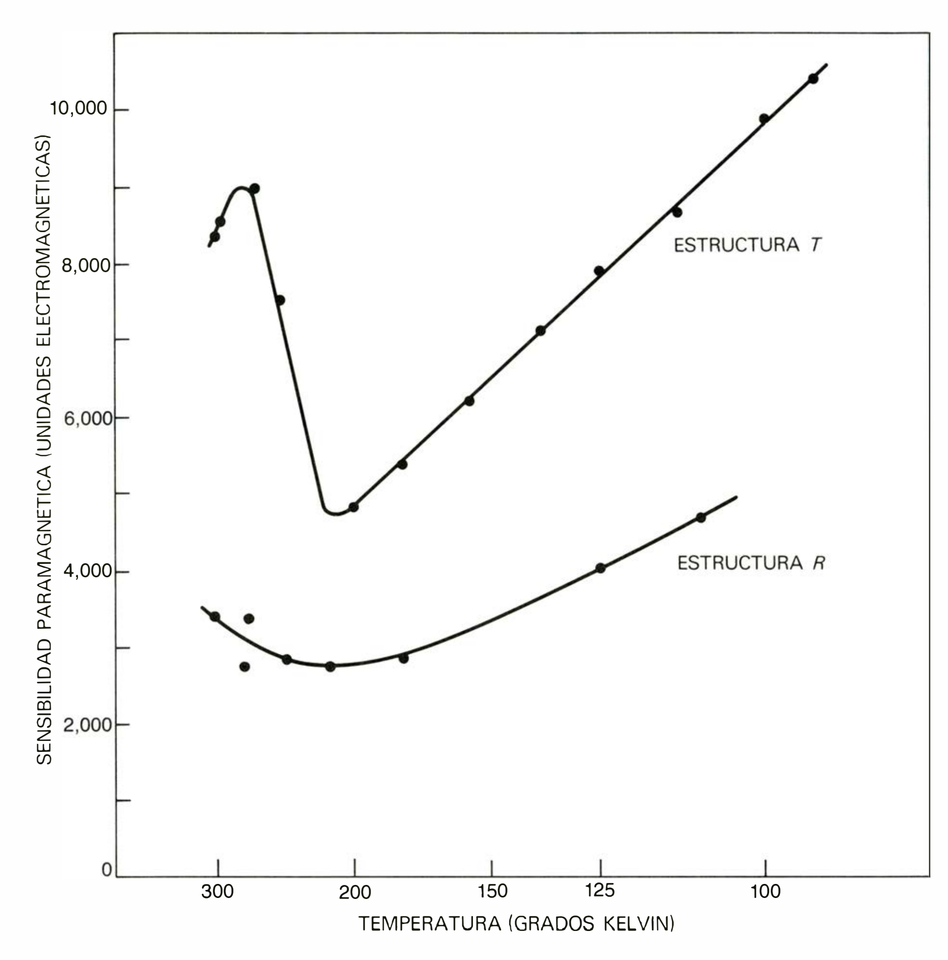
Universidad de Harvard, mi propio colega Joyce Baldwin y Cyrus Chothia, del Colegio Universitario de Londres, habían tratado de desenredar el conjunto de palancas atómicas que generan la tensión en la estructura *T* y liberan ésta en la transición hacia la estructura *R*. Ellos han demolido algunas de mis ideas iniciales y han elaborado otras.

Todos están de acuerdo en que la estructura *T* ejerce una tensión pequeña o nula en el grupo hemo desoxigenado, y que la tensión aparece solamente cuando el hierro trata de trasladarse hacia el plano de la porfirina en combinación con el oxígeno, de forma parecida al muelle de una puerta de cierre automático que está relajado cuando ésta se halla cerrada, pero que ejerce una tensión que va incrementándose progresivamente a medida que la puerta se abre. James P. Collman, de la Universidad de Stanford, ha sugerido en este sentido que debería hablarse de sujeción en vez de tensión. La sujeción podría estar generada por una orientación dirigida más hacia la histidina proximal con respecto a la porfirina, lo cual haría que uno de los átomos de carbono de la histidina estuviera más próximo a un átomo de nitrógeno de la porfirina. La repulsión entre estos dos átomos evitaría que la histidina se acercara al anillo porfirínico.

En la transición a la estructura *R*, un desplazamiento y rotación del hemo en relación con la hélice *F* endereza la histidina, de manera que ésta y el átomo de hierro se pueden trasladar hacia la porfirina sin sujeción.

En las subunidades beta, el movimiento del grupo hemo con respecto a la hélice *E*, que transporta la valina distal y la histidina distal, puede ser más importante. En la estructura *T*, la valina bloquea el lugar de combinación del oxígeno, pero después del desplazamiento a la estructura *R*, el lugar permanece sin cubrir. No sabemos todavía hasta qué punto esta ocupación y desocupación contribuye a la energía libre de la interacción hemo-hemo.

Todos estos mecanismos están de acuerdo con mis ideas iniciales, aunque mi sugerencia de que el movimiento de la histidina proximal se transmite a los puentes salinos por la compresión de la penúltima tirosina hacia fuera de su hendidura era demasiado simplista. En vez de ello, el enlace de hidrógeno que mantiene la tirosina en su lugar puede ser estirado, aunque este aflojamiento quizá no resulte suficiente para romper los puentes salinos. Estos pueden ser desprendidos posteriormente por pequeñas perturbaciones de los enlaces entre



CAMBIOS MAGNETICOS registrados al pasar la metahemoglobina azida de la carpa, de la estructura *R* a la *T*. La sensibilidad paramagnética está representada en el eje vertical y la temperatura absoluta en el eje horizontal. El paramagnetismo de los átomos de hierro es mayor en la estructura *T* que en la estructura *R*, para todas las temperaturas. En la estructura *R* a temperaturas bajas, todos los átomos de hierro se encuentran en un estado paramagnético débil y la sensibilidad desciende al subir la temperatura. A unos 200 grados Kelvin, los átomos de hierro empiezan a oscilar entre un estado paramagnético débil y otro fuerte, de manera que la susceptibilidad aumenta al hacerlo la temperatura. En la estructura *T* a temperatura baja, una mezcla al azar de átomos de hierro en estado paramagnético fuerte y débil se congela. A unos 250 grados K, la fracción de átomos de hierro fuertemente paramagnéticos empieza a aumentar de un modo brusco, para descender de nuevo a temperaturas más altas por razones que aún no están claras. El equivalente de energía libre de la tensión en la estructura *T* se calcula por la diferencia en altura entre las dos curvas. La susceptibilidad magnética mide la fuerza que se ejerce en un equivalente gramo de hierro (55,8 gr) por un campo magnético de un gauss (unidad CGS que equivale a un maxwell por centímetro cuadrado).

las subunidades cuyo análisis ha sido elusivo hasta ahora.

Una de las características más extrañas tanto de la estructura *T* como de la *R* es la ausencia de cualquier entrada a la cavidad del hemo cuya apertura permitiera el paso de una molécula de oxígeno. La histidina o cualquier otro grupo deben balancearse fuera de la entrada, pero no sabemos cómo ocurre esto, puesto que nuestros análisis de rayos X reflejan estructuras estáticas, que nos permiten solamente conjeturar la dinámica de la molécula.

John J. Hopfield, de la Universidad de Princeton, dijo una vez que la hemoglobina desempeña el mismo papel en bioquímica que el átomo de hidrógeno en física, porque sirve de piedra de toque para nuevas teorías y técnicas experimentales. La hemoglobina es el prototipo de molécula proteica que cambia su estructura en respuesta a un estímulo químico. Los científicos, por tanto, continuarán explorando su comportamiento polifacético. El mecanismo que yo he esbozado aquí necesitará estudios adicionales antes de que puedan explicarse todas las observaciones incluidas, pero yo estoy contento de que sus características fundamentales hayan salido airosas de las pruebas experimentales, y de que éstas se ajusten razonablemente bien a las propiedades fisiológicas de la hemoglobina. No he mencionado aquí que esto también explica los síntomas de los pacientes que han heredado hemoglobinas anormales, porque es otra historia. Espero que la comprensión de la estructura y mecanismo de la molécula de hemoglobina sirva para aliviar esos síntomas y para explicar el comportamiento de sistemas biológicos que encierren mayor complejidad.

Ojos animales donde la imagen se forma mediante espejos

Hoy se construyen espejos cuyas superficies reflectoras están formadas por una multitud de capas finas y no por una placa metálica. Existen varios tipos de animales cuyos ojos están basados en el mismo sistema

Michael F. Land

El descubrimiento de que la imagen que se visualiza en el ojo humano se forma mediante una lente se remonta a finales de la Edad Media, época a partir de la cual se empezaron a usar lentes de vidrio como gafas correctoras. En los siglos posteriores, los espejos también llegaron a ser componentes importantes de los instrumentos ópticos tales como el telescopio astronómico, pero nadie podía albergar la sospecha de que las superficies reflectoras pudieran constituir el mecanismo óptico básico de los ojos de determinados animales. En estos últimos años, se ha puesto de manifiesto que existen varios ejemplos sorprendentes de ojos cuyo sistema óptico se basa en espejos: entre ellos hallamos los ojos simples de las veneras o conchas de peregrino y los ojos compuestos de los langostinos, los cangrejos de río y las langostas.

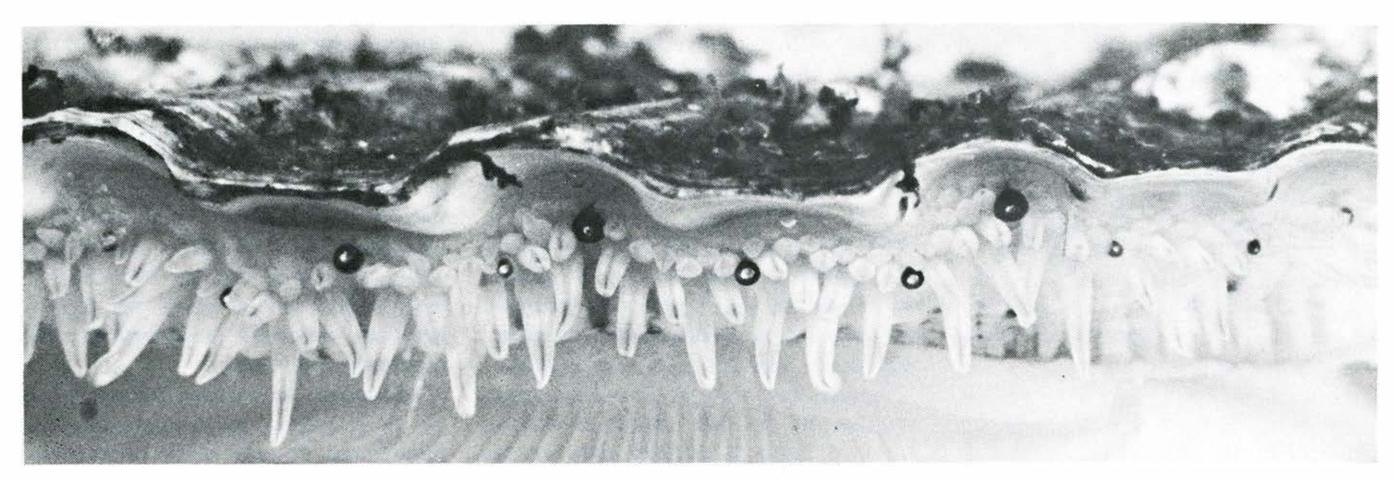
No se puede decir que estos animales sean particularmente exóticos; es más, la anatomía de sus ojos ha sido descrita en numerosas ocasiones. Uno se llega a preguntar, entonces, cómo es posible que se hubiera pasado por alto el papel de los espejos en la formación de la imagen visual, a pesar de la multitud de investigaciones anatómicas y ópticas que se dedicaron al estudio de los ojos de los animales invertebrados. Quizá la respuesta esté en que hasta hace relativamente poco tiempo se consideraba como una imposibilidad biológica la existencia de espejos de calidad óptica, no alcanzando dicha categoría las porciones brillantes que podían hallarse en los tejidos. Estas ideas tenían una lógica sencilla: la superficie de un espejo es metal pulido y los seres vivos no fabrican superficies metálicas. No obstante, desde la década de los cuarenta han variado las técnicas utilizadas para elaborar espejos de alta calidad. Actualmente, en lugar de consistir en una capa única de plata o aluminio, los espejos se fabrican mediante pilas estratificadas formadas por un gran número de películas muy finas cuyo índice de refracción es alternativamente elevado y bajo. ¡Ahora se ha comprobado que éste es precisamente el principio en que se han basado desde siempre los espejos de los organismos vivos! Es evidente que el progreso tecnológico ha permitido eliminar el importante obstáculo de la reticencia mental al concepto de los espejos animales.

El funcionamiento de los espejos formados por multicapas es fácil de entender. Las burbujas de jabón constituyen un ejemplo sencillo de película delgada, cuya propiedad reflectante es el resultado de la interferencia de las ondas luminosas reflejadas desde las superficies externa e interna de la película. Si dos frentes de onda quedan reflejados en fase a partir de las dos interfases, la reflectancia resulta elevada, pero si los frentes de onda no están en fase, las crestas de las ondas anulan los valles y la reflectancia queda reducida a cero. Para que los frentes de onda queden en fase al reflejarse, el espesor óptico de la película (es decir, su espesor real multiplicado por el índice de refracción) debe ser igual a una cuarta parte de la longitud de onda de la luz incidente. La explicación de este hecho es que las ondas que quedan reflejadas en la superficie superior sufren un cambio de fase de media longitud de onda, mientras que las ondas reflejadas en la superficie inferior no sufren cambio de fase alguno; además, las ondas reflejadas desde la superficie interna atraviesan la película por dos veces en direcciones opuestas.

Con una sola película de cuarto de longitud de onda, queda reflejado un 8 por ciento de la luz incidente, pero se pueden conseguir espejos que reflejen la totalidad o la casi totalidad de la luz incidente si se apilan varias películas de cuarto de longitud de onda, dejando entre ellas espacios de un cuarto de longitud de onda. Este dispositivo garantiza que, si se ilumina con luz de la longitud de onda adecuada, ésta queda reflejada en fase a partir de todas las interfases de la pila de películas, obteniéndose como resultado una reflectancia superior a la de una placa metálica pulida, a condición de que el número de películas sea superior a ocho.

Los espejos formados por multicapas se pueden construir a partir de materiales diversos, siempre y cuando exista una gran diferencia en el índice de refracción de las capas adyacentes de la pila. Los que se venden en los comercios suelen estar constituidos por películas alternas de fluoruro de magnesio (índice de refracción 1,36) y de sulfuro de cinc (2,40), depositadas a partir de los vapores respectivos. En los materiales biológicos, es imposible una diferencia tan amplia en los índices de refracción, pero una combinación biológica corriente, hallada tanto en las escamas de los peces como en el ojo de la venera, es la alternancia de citoplasma (1,34) y cristales de guanina (1,83). Los espejos de los insectos a menudo están formados por capas alternas de aire (1,00) y el material polisacárido denominado quitina (1,56).

Los espejos de multicapa poseen una serie de propiedades interesantes de las que carecen los espejos metálicos. Debido a que tan sólo una longitud de onda resulta ser cuatro veces el espesor óptico de las películas de la pila, la luz reflejada es de color; un ejemplo biológico especialmente sorprendente es el de las escamas de las alas de las mariposas del género *Morpho*, que son de un color azul irisado. También se pueden conse-



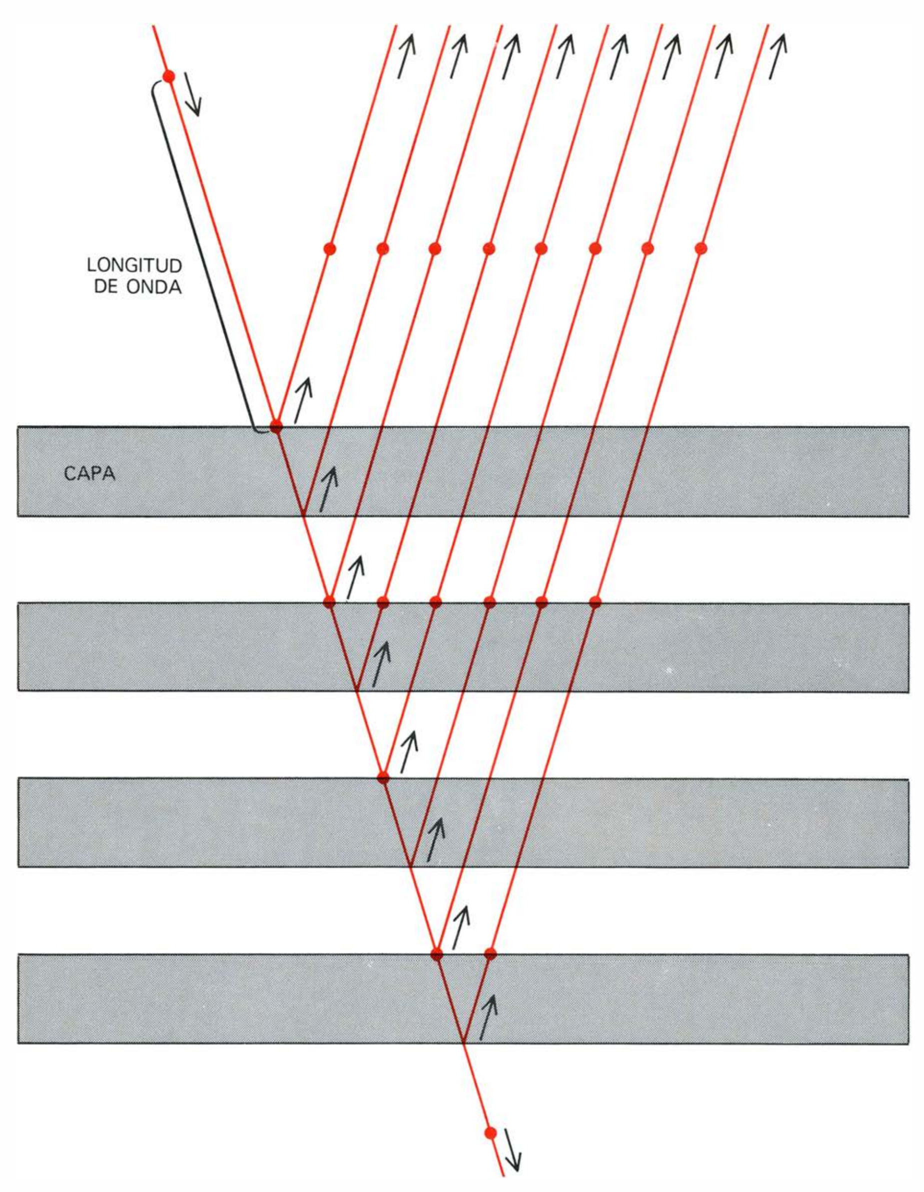
OJOS DE LA VENERA *Pecten*. Se trata de los globulillos que se observan en esta fotografía del molusco donde aparece el manto que queda expuesto al cerrarse las dos valvas, o conchas, dejando una pequeña hendidura. Una de las dos valvas se halla en la parte superior. Cada uno de los ojos, cuyo diámetro es de aproximadamente un milímetro, consiste en un espejo semiesférico diminuto. Los ojos son capaces de detectar el movimiento de una sombra o de un objeto de borde obscuro dentro del campo de visión del animal. Y así es como descubre la aproximación de un depredador. Los ojos intervienen también en las reacciones de fototaxismo: la venera puede migrar hacia zonas más obscuras o más iluminadas del medio circundante.

LA IMAGEN que vemos aquí en el interior del ojo de la venera es la de una rejilla cuadrada situada entre el ojo del animal y el objetivo del microscopio. La rejilla queda enfocada por el espejo situado en la parte trasera del ojo. El ojo posee una lente que corrige la aberración de esfericidad del espejo.

guir espejos de luz blanca a partir del principio de multicapa, tanto en la naturaleza como en el laboratorio: basta con variar el espesor de las películas a las distintas profundidades de la pila o amontonar tres pilas distintas, que reflejen respectivamente la luz azul, la verde y la roja. Algunos peces utilizan este sistema para presentar una apariencia blanca, a pesar de que sus escamas multiestratificadas individuales son fuertemente coloreadas.

El color de un espejo de multicapa varía acercándose a las longitudes de onda más cortas (azul y violeta) al aumentar el ángulo de incidencia de la luz. Este fenómeno es el resultado del cambio en la longitud del camino óptico de la luz en las capas de la pila. Por otra parte, los espejos de multicapa se diferencian de los filtros en que no absorben luz; las longitudes de onda que no son reflejadas se transmiten en forma del color complementario. Este tipo de espejos se emplea en las cámaras de televisión en color con el fin de separar, sin pérdidas de luz, los tres colores primarios.

La venera común *Pecten* tiene aproximadamente 60 ojos, azules o marrón obscuro (según la especie), cada uno de un milímetro de diámetro, espaciados de forma regular a lo largo del manto del molusco, que se halla expuesto a la luz por una hendidura delgada que queda entre las dos valvas o conchas. Los ojos sirven para detectar el movimiento: si a una venera de gran tamaño situada en un acuario o en el fondo marino se le acerca un posible depredador, reacciona enterrándose en la arena; las especies menores responden alejándose por palmoteo de sus valvas.



UNA SERIE DE CAPAS MULTIPLES formadas por materiales de índice de refracción alternadamente alto y bajo puede resultar en una superficie reflectora dotada de un elevado grado de reflectancia. Esta técnica se utiliza hoy en día corrientemente en la fabricación de espejos de alta calidad para telescopios y otros instrumentos; resulta que también es el método que emplean los organismos vivos que llevan espejos incorporados. Cada capa de la pila debe tener un espesor que sea un cuarto de la longitud de onda de la luz incidente, con el fin de que las ondas reflejadas a partir de todas las interfases de la pila se encuentren en fase. Debido a que una sola longitud de onda es la que se halla reflejada en fase, el espejo es fuertemente coloreado. El color varía hacia el extremo azul del espectro al aumentar el ángulo de incidencia de la luz que ilumina la multicapa y al disminuir, por tanto, la longitud del camino óptico de los rayos a través de las capas de la pila. (Dibujo de Tom Prentiss.)

Observado en sección transversal, el ojo de la venera no se diferencia, aparentemente, de los ojos de los animales vertebrados, y tal vez ésta sea la razón por la que su sistema óptico particular ha pasado inadvertido durante tanto tiempo. Presenta una gran lente transparente situada por delante de una retina de dos capas, constituidas cada una por unas 5000 células fotorreceptoras. Por detrás de la retina, existe un reflector semiesférico, seguido por una capa de pigmento marrón. Hay una característica de la anatomía del ojo de la venera que hace sospechar que su sistema óptico no es el convencional. En los ojos de los vertebrados, siempre existe un espacio por detrás de la lente, el cual permite que la imagen pueda quedar enfocada sobre la retina. Incluso en el ojo de los peces, cuya lente suele tener una distancia focal muy corta en relación con su diámetro, existe un espacio entre la lente y la retina cuya anchura es de unas dos veces el radio de la lente. En cambio, en el ojo de la venera, la parte posterior de la lente se halla en contacto con la retina, de modo que la imagen enfocada se situaría más allá de la parte posterior del ojo. Lo que está claro es que o bien el ojo de la venera no forma una imagen útil para el animal, o bien la imagen se forma de un modo distinto al descrito.

Si uno deja penetrar su mirada directamente en el ojo de una venera fresca por medio de un microscopio de disección, observa una imagen más que sorprendente: ve una pequeña y brillante imagen de sí mismo, invertida y algo distorsionada, que parece estar observando atentamente a través del microscopio. La imagen parece situarse en algún lugar en el interior del ojo de la venera, entre el espejo y la lente. Esta observación no constituye ninguna novedad: el zoólogo americano William Patten ya la había realizado hace casi un siglo, llegando incluso a pintar líneas sobre el objetivo de su microscopio con el fin de examinar la imagen con más claridad. Patten llegó a la conclusión errónea de que la lente del ojo de la venera era responsable de la formación de la

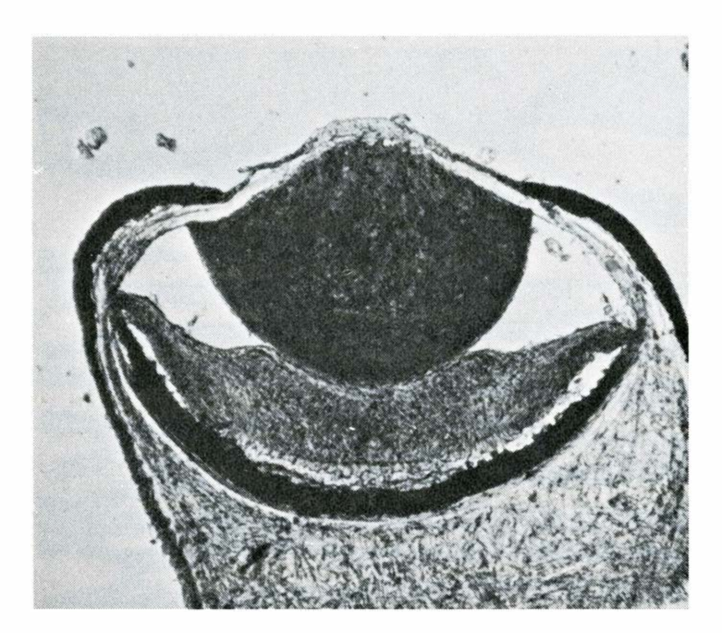
imagen y que el espejo tan sólo se limitaba a reflejar otra vez la imagen hacia el exterior, como si de la reflexión del ojo de un gato se tratara. Pero éste no es el caso, por los mismos motivos por los que no se puede ver una imagen en el ojo de un gato: si uno puede ver una imagen a través de la misma lente que la formó, lógicamente la imagen se halla situada fuera del ojo. Sin embargo, tanto Patten como yo hemos observado que la imagen se halla en el interior del ojo de la venera.

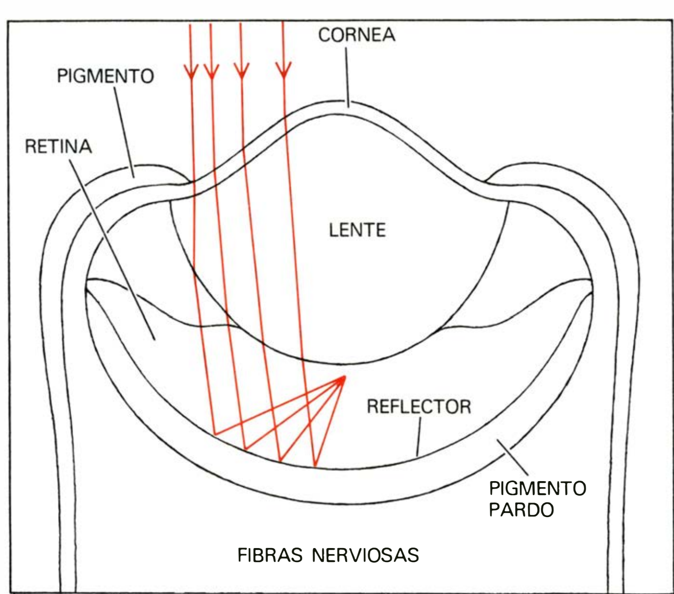
La solución al problema se hizo patente cuando me di cuenta de que es pequeño el efecto de la lente. El espejo es una semiesfera casi perfecta y, al igual que cualquier otro espejo cóncavo, forma una imagen invertida de los objetos alejados, situada a medio camino entre el centro de curvatura del espejo y el propio espejo. El reflector del ojo de la venera tiene un radio de curvatura de 410 micras y, por tanto, la distancia focal del espejo sería de 205 micras. Algunas mediciones que yo había efectuado cuando aún estudiaba la licenciatura en el University College London pusieron en evidencia que, efectivamente, la posición de la imagen invertida correspondía al valor predicho, aunque la distancia focal del espejo era menor en 50 micras que la que podría deducirse del concepto de un "reflector puro". Esta diferencia puede quedar explicada por el poder de refracción de la lente del ojo de la venera. Aunque la lente tiene una distancia focal de 1,5 milímetros y, aislada, formaría una imagen situada muy por detrás de la retina, su efecto real es el de hacer converger ligeramente los rayos de la luz incidente antes de que alcancen el espejo, acortando así de alguna manera la distancia que hay entre el espejo y la imagen.

La forma de la superficie anterior de la lente es un tanto extraña: es convexa en el centro y cóncava en los bordes, por lo que parece que su función no sería primordialmente la de hacer converger los rayos, sino la de corregir algún defecto óptico del espejo. Cualquier espejo semiesférico está sujeto a la aberración esférica, ya que los rayos reflejados por la parte interior del espejo convergen hacia un foco situado más hacia fuera que el foco de los rayos reflejados en las partes más externas del mismo. En los telescopios astronómicos del tipo de Schmidt, se corrige la aberración esférica de un espejo semiesférico mediante una lente fina de poco poder de refracción, denominada placa correctora, que se sitúa en el centro de curva-



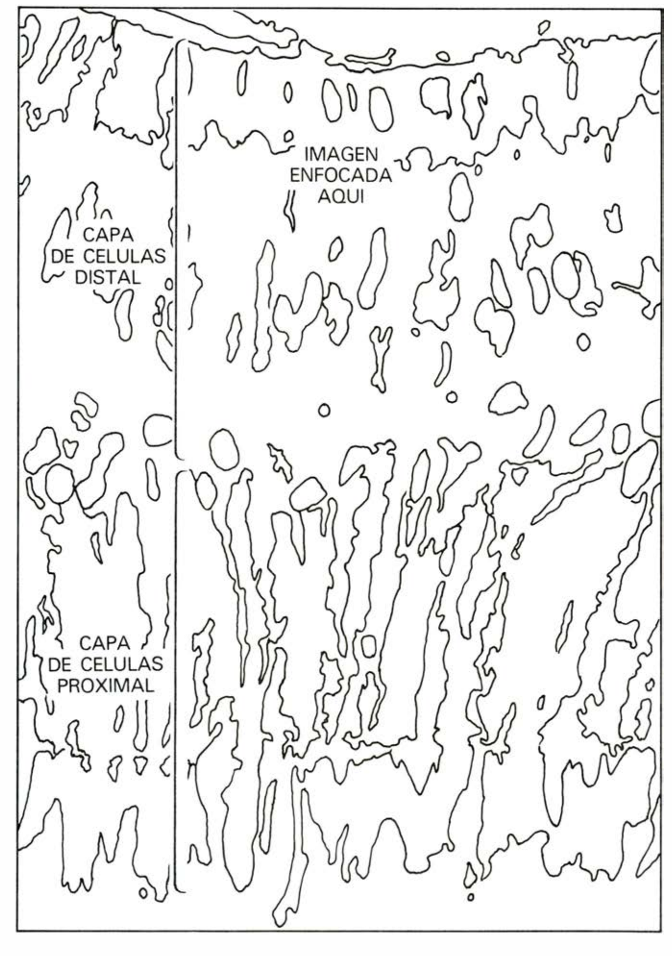
ESPEJO DEL OJO DE LA VENERA, que está construido por capas alternas de citoplasma (índice de refracción 1,34) y de cristales de guanina (1,83), como se puede ver en esta microelectrografía de una sección transversal del espejo. (Los agujeros aplanados se hallaban previamente ocupados por los cristales de guanina, los cuales fueron eliminados en el proceso de fijación.) El reflector tiene un grosor de seis micras. Esta micrografía fue realizada por Vernon Barber, del University College London.





SECCION TRANSVERSAL del ojo de la venera. Induce a pensar que su sistema óptico no es el corriente. A diferencia de los ojos con lente de los animales vertebrados, que poseen una zona clara entre la lente y la retina, zona que permite que los rayos converjan en un foco, la lente del ojo de la venera está en contacto directo con una retina en forma de "croissant". Por detrás de la retina, se encuentra un espejo sumamente delgado (no visible con este aumento) seguido por una capa fina de pigmento obscuro. El diagrama de la derecha representa el camino óptico de los rayos luminosos en el interior del ojo. Los rayos, que sufren tan solo una muy ligera refracción al cruzar la lente, atraviesan la retina alcanzando el reflector semiesférico, el cual los devuelve a la retina, enfocando la imagen a nivel de las células fotorreceptoras distales que ocupan la capa superior de la retina. Debido a que la luz pasa una primera vez por la retina antes de quedar detectada la imagen, el ojo no presenta una agudeza satisfactoria en cuanto al contraste.





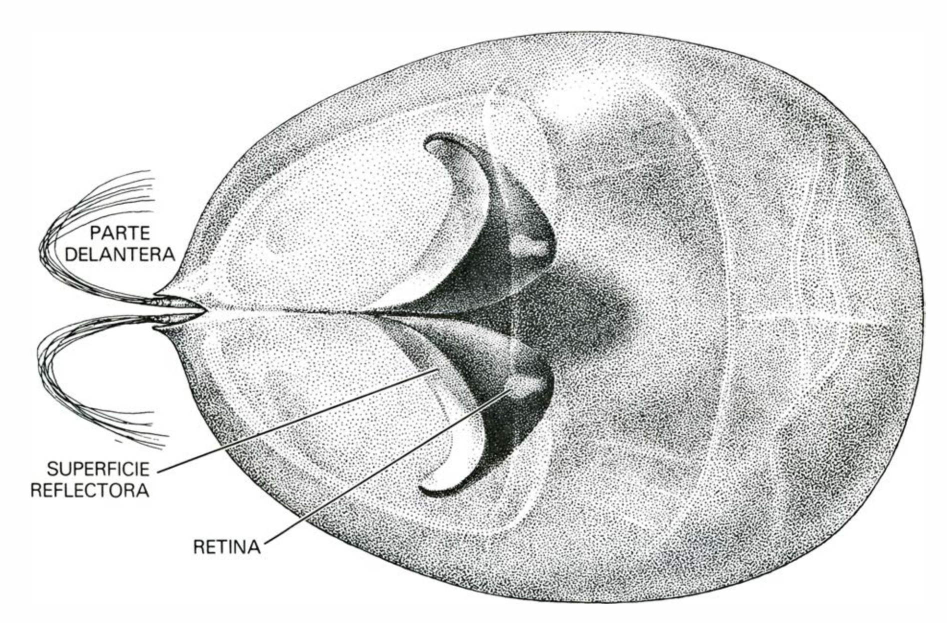
CAPAS DE LA RETINA del ojo de la venera, que contienen las células detectoras de la luz, según aparecen en esta sección transversal. En la capa de células distales situada inmediatamente detrás de la lente, unas 5000 células fotorreceptoras reciben una imagen enfocada procedente del reflector. Estas células responden al estímulo de un borde obscuro que atraviesa el campo de visión; el movimiento es captado debido al amortiguamiento secuencial de la intensidad de la luz en los fotorreceptores adyacentes. La capa de células proximales situada en la parte inferior de la retina recibe una imagen desenfocada procedente del reflector. Estas células responden selectivamente a un aumento global, de un determinado tanto por ciento, de la intensidad de iluminación del campo de visión entero. Están implicadas en el comportamiento fototáctico y no en la discriminación de imágenes.

tura del espejo. Resulta admirable que la lente del ojo de la venera parezca llevar a cabo una función muy semejante a la de la placa correctora en el telescopio de Schmidt.

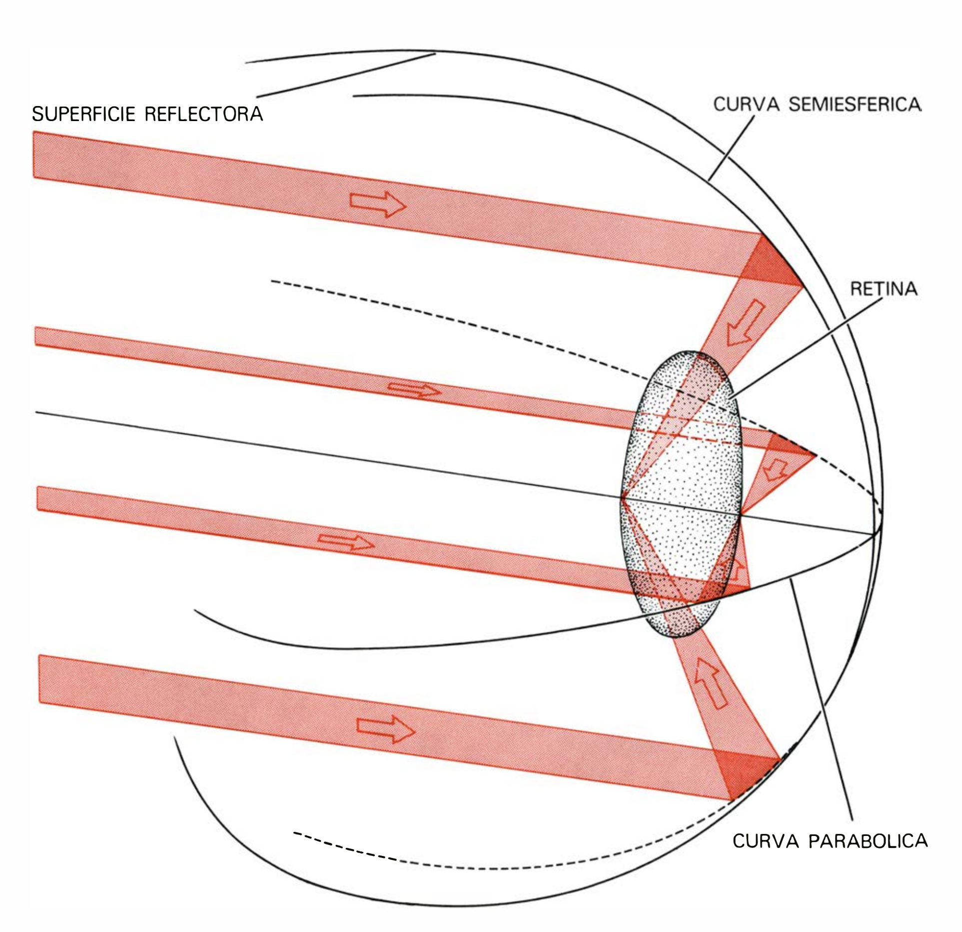
La imagen formada por el reflector del ojo de la venera queda enfocada en la parte de la retina situada inmediatamente detrás de la superficie posterior de la lente. La pregunta obligada que se plantea es si existen o no en esta parte de la retina células especializadas para recibir luz. En realidad, la retina presenta dos conjuntos de células sensibles a la luz: una capa de células distal, situada inmediatamente detrás de la superficie posterior de la lente, y una capa de células proximal, situada directamente enfrente del espejo. La imagen formada por el espejo tiene su foco situado en los fotorreceptores distales, mientras que los fotorreceptores proximales reciben una imagen desenfocada.

Ya en el año 1938, H. K. Hartline, del Cornell University Medical College, efectuó unos registros electrofisiológicos de los fotorreceptores distales de la venera y comprobó que estas células responden con un impulso nervioso tan sólo cuando el estímulo es la privación de la iluminación (es decir, una respuesta de tipo "off"). Por tanto, estas células serían activadas por la imagen del extremo anterior de un objeto obscuro que cruzara el campo de visión. Si el objeto es más claro que el medio circundante, el estímulo lo constituiría el extremo posterior. Sin embargo, los objetos de color claro no suelen ser estímulos tan eficaces, debido a que los fotorreceptores distales han de estar iluminados durante varios segundos de forma continua antes de poder generar una respuesta de tipo "off".

He continuado investigando sobre la función de la retina de la venera. En uno de mis experimentos, se registraba la actividad electrofisiológica de los fotorreceptores distales mientras se hacía pasar por delante del ojo de la venera un motivo de franjas negras y blancas alternadas, en que cada franja tenía una anchura de 16 grados. Cuando se desplazaba el motivo una distancia correspondiente a la anchura de una franja (de modo que las áreas que habían sido iluminadas estuvieran ahora obscuras y viceversa) sin variar el nivel de la iluminación, los fotorreceptores distales emitían descargas de forma masiva y continuada. Cuando el motivo se desplazaba una anchura de dos franjas (de modo que cualquier punto de la retina fuera brevemente iluminado o privado de su iluminación antes de volver a restaurarse la imagen inicial), tan sólo se dio



CRUSTACEO DE LAS GRANDES PROFUNDIDADES MARINAS ***Gigantocypris.*** **Posee grandes ojos reflectores que le permiten concentrar la luz sumamente débil disponible a la profundidad oceánica de 1000 metros (procedente principalmente de los órganos luminosos de los crustáceos y peces luminosos). El animal mide aproximadamente un centímetro, tamaño que resulta grande en relación con la escala de sus parientes planctónicos: la cabeza viene a representar casi la mitad del cuerpo. Ambos ojos reflectores están cubiertos por ventanas transparentes practicadas en el caparazón naranja que encierra totalmente al animal. Sir Alister Hardy comparó estos ojos con "los faros de un gran coche". Hardy fue el primero en especular con la posibilidad de que estos espejos sirvieran para enfocar la luz.**



OPTICA DEL OJO REFLECTOR de ***Gigantocypris,*** **bastante compleja. En la dimensión horizontal, el reflector situado en la parte posterior del ojo es parabólico y el punto focal de la parábola se encuentra a una corta distancia delante del espejo. En la dimensión vertical, en cambio, el reflector es semiesférico y tiene su punto focal situado más lejos del espejo, de tal suerte que el conjunto del espejo enfoca la luz a lo largo de una línea que se extiende entre el foco del reflector parabólico y el del reflector semiesférico. Esta línea focal coincide con la localización de unas largas células fotorreceptoras situadas en una retina con forma asalchichada, que se halla delante mismo del espejo. Aunque la agudeza de este ojo reflector es baja, forma una imagen que resulta 17 veces más intensa que la imagen que se hubiera formado en el ojo de lente que es típico de los peces (cuyo número** ***f*** **vale 1,25).**

una emisión breve de impulsos en los fotorreceptores distales. Estos resultados sugieren que la respuesta de evitación de la venera está desencadenada por un amortiguamiento secuencial de la intensidad de la luz recibida en las células fotorreceptoras adyacentes de la capa distal de la retina, y no por la detección del movimiento en sí.

¿Cuál puede ser, entonces, el papel de los fotorreceptores de la capa proximal, que reciben del espejo una imagen desenfocada? Hartline registró también la actividad de estas células y encontró que emitían impulsos en respuesta a la iluminación (una respuesta de tipo "on", en actividad), pero tan sólo cuando el objeto que entra en escena es tan grande y luminoso que provoca un aumento de intensidad de un determinado tanto por ciento en la totalidad del campo de visión. A pesar del bajo poder de resolución de estas células en la detección del movimiento, es posible que desempeñen algún otro papel importante. Se sabe que algunas de las especies de tamaño menor se desplazan hacia las regiones más claras o más obscuras del medio que las rodea, y es más probable que este comportamiento fototáctico implique la participación de los fotorreceptores proximales, cuya respuesta suele ser mantenida, y no la de los fotorreceptores distales, cuya respuesta es siempre transitoria. Por tanto, los fotorreceptores proximales son responsables de recibir una información a plazo relativamente largo acerca de los niveles de iluminación ambientales.

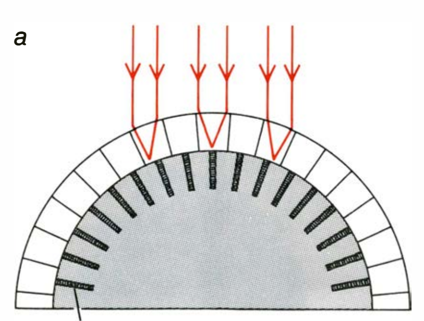
Lo que sigue siendo un misterio es por qué el ojo de la venera tiene un número de fotorreceptores proximales tan elevado. Con 60 ojos y 5000 fotorreceptores proximales por ojo, el molusco viene a tener 300.000 células para realizar un trabajo que aproximadamente una docena de ellas podrían hacer adecuadamente. Tampoco los fotorreceptores distales se utilizan de modo rentable: cada ojo tiene un campo de visión de 100 grados aproximadamente, ¡y no se necesitan 60 campos de 100 grados para cubrir 360 grados!

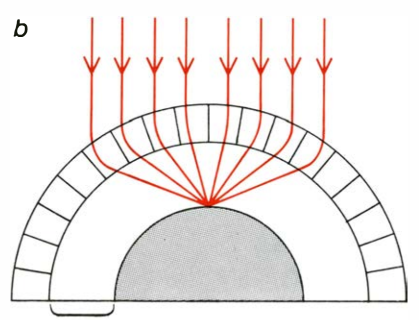
La búsqueda de otros ojos en el reino animal que funcionaran según los mismos principios ópticos que los de *Pecten* ha resultado aportar, sorprendentemente, muy pocos casos de estos ojos. Varios berberechos y unos cuantos crustáceos copépodos tienen ojos que llevan espejos incorporados, pero los ojos son tan diminutos que no es posible ni tendría sentido siquiera intentar determinar si es el espejo el que forma la imagen o si actúa simplemente doblando la longitud del camino óptico.

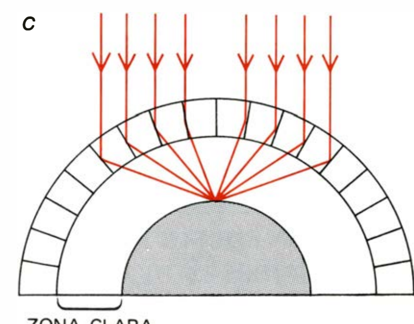
Sin embargo, existe un invertebrado que tiene grandes ojos de tipo reflectante: es el crustáceo *Gigantocypris* de los fondos marinos, que, al tener un centímetro de longitud, resulta realmente gigantesco en la escala de sus parientes componentes del plancton. Sir Alister Hardy, de la Universidad de Oxford, describió el animal en 1956, haciendo la observación de que "el par de ojos presentan en su parte posterior inmensos reflectores que parecen metalizados, dándoles el aspecto de los faros de un gran coche; se proyectan hacia el exterior a través de unas ventanas como de vidrio practicadas en el caparazón que es naranja, excepto en estos puntos; sin duda alguna, los espejos cóncavos posteriores sustituyen a una lente que estaría en la parte anterior". A mi juicio, ésta fue la primera ocasión en la que apareció en la literatura biológica una hipótesis seria en relación con sistemas biológicos ópticos basados en espejos. Anteriormente a Hardy, otros investigadores habían pretendido explicar estos ojos aduciendo que el caparazón de forma redondeada en el contorno de los ojos podía actuar como una lente; pero debido a que el caparazón es delgado, tiene caras paralelas y se halla en contacto con agua a ambos lados, dicha hipótesis resulta imposible. El único sistema óptico que pueda darse es el propuesto por Hardy.

Hace poco se me ofreció la posibilidad de estudiar, en mi laboratorio de la Universidad de Sussex, ejemplares vivos y conservados del género *Gigantocypris*, gracias a la generosidad del Instituto de Oceanografía de Wormley, y actualmente estoy en condiciones de imprimir un giro interesante a las especulaciones de Hardy. Cuando se miran estos ojos desde arriba, tienen una forma claramente parabólica, estando situado el punto focal de la parábola a una distancia muy corta de la parte anterior del espejo (unas 350 micras). Sin embargo, si se miran lateralmente, la configuración del espejo parece semiesférica, como la de *Pecten*, con una distancia focal mucho mayor (aproximadamente 670 micras).

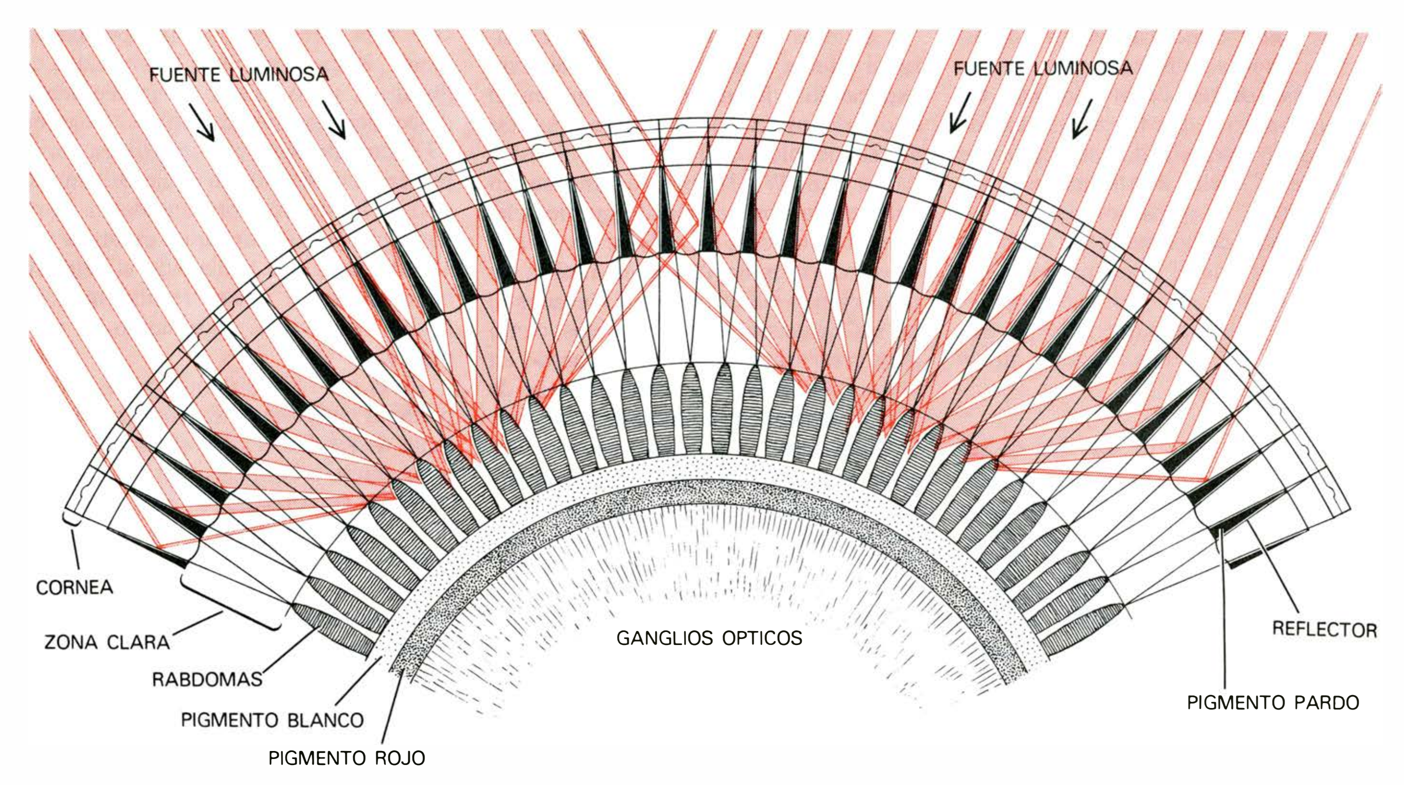
El significado de esta forma tan peculiar se puede deducir si imaginamos que penetra directamente en el ojo un hilo de luz horizontal formado por una hendidura. Si giráramos lentamente el hilo de luz hasta que estuviera orientado verticalmente, a medida que fuera girando su imagen se iría moviendo desde el







SON TRES TIPOS DIFERENTES de ojo compuesto los que se observan en los insectos diurnos, en los nocturnos y en los crustáceos del suborden de los Macruros (langostas, langostinos y cangrejos de río). Los ojos de los insectos diurnos (*a*) están formados por una ordenación hexagonal de pequeñas lentes, cada una de las cuales tiene su propio conjunto de células fotorreceptoras. Los ojos de los insectos nocturnos (*b*) presentan también una malla de lentillas hexagonales, pero (debido a que cada lentilla tiene un gradiente radial del índice de refracción) consiguen desviar la luz de una forma continua de modo a hacer converger los múltiples rayos que penetran en el conjunto del ojo en un solo punto de la retina, aumentando así la intensidad de la imagen en condiciones nocturnas. Los ojos de superposición de los insectos nocturnos presentan una zona clara entre la ordenación de lentillas y la retina, espacio que sirve para que converjan los rayos. Los ojos de los crustáceos macruros (*c*) están formados por una malla cuadrada de cajitas forradas de espejos que actúan superponiendo los rayos luminosos sobre un solo punto de la retina. El que estos ojos superpusieran la luz y presentaran una zona clara hizo que los anatomistas postularan, erróneamente, que funcionaban siguiendo el mismo mecanismo óptico que los ojos de los insectos nocturnos. Hasta 1975 no se cayó en la cuenta de que en estos crustáceos la imagen se forma gracias a una ordenación de espejos.



SUPERPOSICION DE LOS RAYOS en el ojo de los crustáceos macruros, aquí esquematizada mostrando el efecto de la iluminación con luz procedente de dos direcciones diferentes. En cada caso, los rayos convergen en un solo punto de la retina, donde se detectan por las células fotorreceptoras.

foco de la superficie parabólica del espejo hacia el foco de su superficie semiesférica. Una posterior rotación de 90 grados haría moverse a la imagen en camino inverso siguiendo la misma línea. Esta geometría poco habitual significa que un conjunto de rayos luminosos paralelos (caso que forzosamente incluye a todos los hilos luminosos posibles) quedaría enfocado a lo largo de una línea perpendicular a la superficie del espejo y que se extiende entre ambos puntos focales.

¿Qué finalidad puede tener esta extraña imagen astigmática? La estructura de la retina de *Gigantocypris* es la que nos da la respuesta. Es un órgano en forma de salchicha que ocupa una región relativamente pequeña en el centro de cada ojo, situada enfrente mismo del espejo. Las células fotorreceptoras son enormes (unas 700 micras de largo y 25 micras de diámetro) y sus largos ejes se hallan orientados formando ángulos rectos con respecto al espejo, de tal manera que se sitúan aproximadamente a lo largo de la línea focal formada por el espejo. Es bastante probable que la calidad de la imagen formada en el ojo sea pobre, pero la combinación de una gran apertura (número *f* aproximadamente 0,3) y de unos fotorreceptores largos garantiza que el ojo llegue a captar una proporción muy elevada de la luz disponible, cuestión importante si se tiene en cuenta que en el océano a 1000 metros de profundidad hay realmente muy poca luz. El número *f* de la lente del ojo de un pez es 1,25, de modo que la ventaja de este determinado tipo de ojo con espejo parecería estribar en que implica una ganancia en la claridad de la imagen con respecto a un ojo de lente típico, representada por un factor de aproximadamente 17.

La escasez de ojos simples basados en espejos, como son los de *Pecten* o de *Gigantocypris*, hace pensar que existe algún punto insatisfactorio en su funcionamiento, que compensa y supera la ventaja que representan en términos de poder de recolección de luz. La explicación más plausible es que estos ojos tienen una capacidad pequeña para resolver motivos de bajo contraste. En el ojo de *Pecten*, la luz atraviesa primero la capa de receptores distales sin enfocarse, antes de volver del espejo formando en ella una imagen enfocada. El efecto de esta disposición es que se reduce a la mitad del contraste de la imagen si se compara con el obtenido en un ojo de lente, de modo que el mundo se visualiza como a través de una cierta neblina.

Los crustáceos del suborden de los Macruros, tales como los langostinos, los cangrejos de río y las langostas, también poseen ojos en los que están incorporados espejos, pero son ojos compuestos, con facetas múltiples, y el principio óptico en que se basan es bastante diferente del que se da en los ojos simples de *Pecten* y *Gigantocypris*. La historia del descubrimiento del ojo compuesto con espejos quizá resulte tan interesante como el propio mecanismo de su funcionamiento óptico. En 1891, Sigmund Exner, de la Universidad de Viena, demostró que se podía fotografiar una imagen derecha en el ojo de algunos insectos nocturnos, como la luciérnaga, si se extirpaba el contenido interno del ojo. También puso de manifiesto que la imagen está formada por superposición, es decir, que los rayos luminosos que penetran por las múltiples facetas del ojo convergen en un solo punto de la retina. Este diseño intensifica en gran manera la imagen en condiciones de baja iluminación, rasgo cuyo significado adaptativo es evidente en los insectos nocturnos. Por otra parte, los ojos compuestos de los insectos diurnos están formados por una ordenación hexagonal de los elementos o lentes, cada uno de los cuales tiene su propio conjunto de células fotorreceptoras.

En el ojo de superposición, las minúsculas lentes deben presentar determinadas propiedades ópticas, que no son las que poseen las lentes simples. Al atravesar el eje de la lente, los rayos luminosos han de ser desviados, debiendo comportarse cada elemento óptico como un telescopio de dos lentes. Exner propuso

que las lentes cilíndricas tenían una estructura interna sofisticada, con un elevado índice de refracción a lo largo del eje del cilindro y que fuera decreciendo de un modo determinado hacia las zonas periféricas. Gracias a este gradiente radial del índice de refracción, las lentillas serían capaces de desviar la luz de forma continua, ya que el gradiente les daría las propiedades —de tipo telescópico— adecuadas y permitiría que todos los elementos del ojo de superposición tuvieran un foco común.

Una característica de los ojos de superposición de los insectos nocturnos es que presentan un espacio transparente entre el conjunto óptico y la retina, espacio cuya presencia es fundamental si los rayos de las múltiples lentillas han de ser llevados hacia un foco común. (La mayoría de los insectos diurnos no presentan esta zona clara, ya que cada una de las lentillas tiene su propio conjunto de fotorreceptores particular.) El hecho de que tanto los ojos de los insectos nocturnos como los de los crustáceos macruros tuvieran estas zonas claras llevó a Exner a postular que compartían el mecanismo óptico: la superposición por lentes cilíndricas.

Las ideas de Exner, tras figurar durante setenta años en los libros de texto explicando la formación de la imagen en el ojo de los artrópodos, fueron puestas en tela de juicio en 1962 por el biofísico danés Jan W. Kuiper, basándose en parte en la observación de que los elementos ópticos de los ojos de los crustáceos macruros no presentan índices de refracción lo suficientemente elevados ni lo suficientemente heterogéneos como para crear un sistema óptico basado en lentes cilíndricas. Las observaciones de Kuiper fueron confirmadas en 1975 por P. Carricaburu, del Museo Nacional de Historia Natural de París, quien encontró, en el ojo del cangrejo de río, un índice de refracción bajo y uniforme, concluyendo que no podía dar lugar a imágenes de superposición.

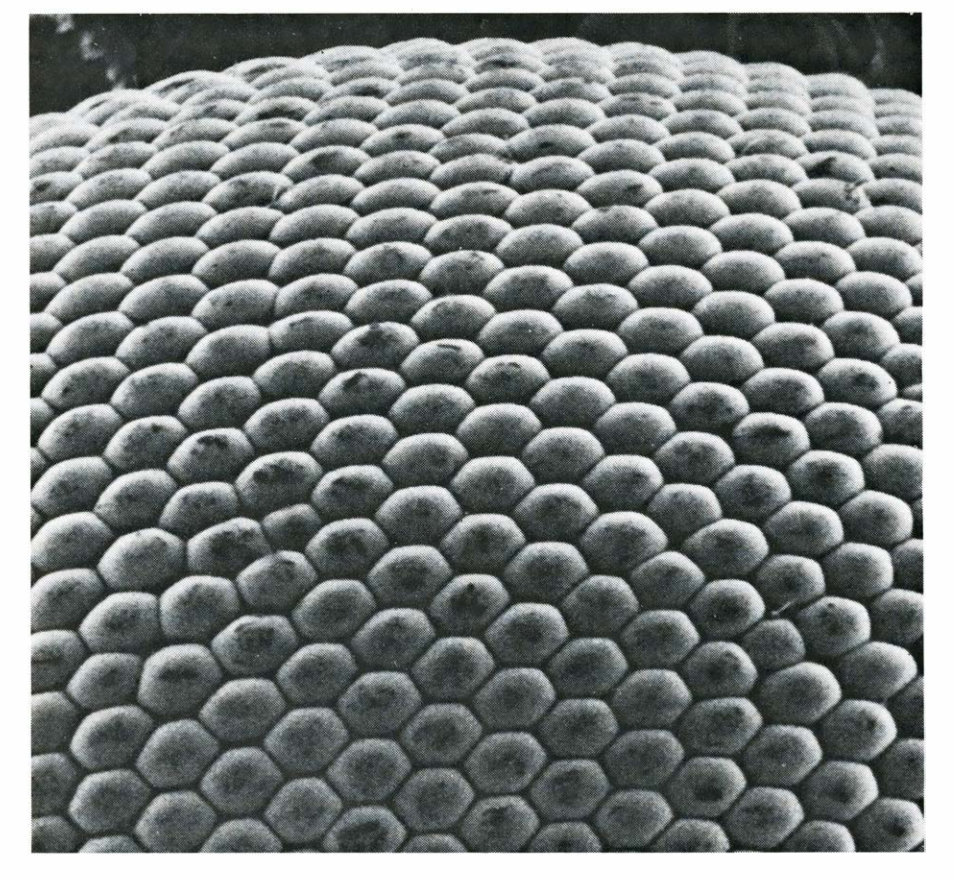
Por tanto, se creaba una situación paradójica: los ojos de los crustáceos macruros se parecen a los ojos de superposición clásicos de los insectos nocturnos; no obstante, los ojos de los crustáceos carecen de las propiedades ópticas adecuadas para que pueda llevarse a cabo la superposición. ¡La solución al dilema era que la superposición no se realiza por lentes cilíndricas sino por diferentes conjuntos de espejos! Esta fue la conclusión a la que llegamos, independientemente, Klaus Vogt, de la Universidad de Stuttgart, y yo mismo. Personalmente, desconocía la publicación de Vogt en 1975 cuando observé por primera vez la presencia de espejos en el ojo del langostino de las grandes profundidades marinas *Oplophorus*; por aquel entonces (en 1976), pensé que había sido el primero en descubrirlo. Luego resultó que, de hecho, yo estaba respaldando la propuesta de Vogt.

En el ojo reflector compuesto, las superficies reflectoras se encuentran delante de los fotorreceptores en lugar de estar detrás y la imagen, en lugar de ser invertida, es derecha. Además, en vez de llevar incorporado un solo espejo, la superficie del ojo se halla dividida formando una matriz cuadrada de pequeñas cajas rectangulares con lados reflectantes, dando lugar en conjunto a una ordenación de facetas reflectantes. La geometría de la ordenación tiene la propiedad de enfocar la luz en la superficie de una esfera cuyo radio es aproximadamente la mitad del radio del propio ojo.

En este caso también resulta un misterio el por qué no se observó anteriormente la presencia de estos espejos. Uno de los factores que puede haber influido es el alto concepto que se tenía de las ideas de Exner mantenido durante mucho tiempo, pero otro motivo probable es que las superficies reflectantes del ojo compuesto no resultan nada evidentes. Los espejos son muy delgados (menos de una micra de espesor) y se destruyen con facilidad en los procesos de fijación, de modo que resulta imposible hallarlos en materiales biológicos preparados de un modo tradicional. Ni siquiera se visualizan en los ojos frescos, porque paralela a la superficie externa de cada espejo hay una franja de pigmento obscuro: sólo se pueden ver las superficies reflectoras verdes si se observa un corte del ojo desde el interior hacia el exterior.

Una vez localizados los espejos, la tarea de elucidar el camino de los rayos y demostrar que era posible la formación de una imagen por superposición no fue difícil. En realidad, uno de los rasgos llamativos de la ordenación de los espejos es que resulta evidente su modo de funcionamiento. En cambio, las lentes cilíndricas de Exner, cuya función básica es la misma, son bastante difíciles de explicar.

Los ojos de tipo reflector compuesto presentan una característica que permite distinguirlos inmediatamente, y sin ningún riesgo de error, de los ojos compuestos de tipo refractor: los ojos reflectores tienen facetas cuadradas mientras



ORGANIZACION HEXAGONAL de las lentillas en el ojo compuesto de tipo refractor de una polilla de la familia de los plutélidos, recogida con nitidez en esta microelectrografía de barrido. Esta ordenación de las lentes es la combinación geométrica más rentable; se da en los ojos de los insectos diurnos.

que los refractores las tienen hexagonales. La forma hexagonal resulta la más económica a la hora de reunir de modo compacto un gran número de unidades, por lo que la aparición de la forma cuadrada es un hecho insólito. Pero, ¿podría estar basado un sistema reflector en una disposición hexagonal? Vogt ha dado una explicación reciente del motivo por el que es tan importante la ordenación en cuadrados. El modo ideal de hacer converger en un solo punto toda la luz procedente de una dirección dada sería con una serie de espejos ordenados formando una red de mallas en forma de bordes de platillo, red que resultaría de cortar transversalmente un conjunto de conos cuyos ápices se hallaran en el centro de curvatura del ojo. Estos espejos presentarían automáticamente la propiedad de que el rayo incidente y el rayo reflejado hacia el foco se encontrarían en el mismo plano. El problema con un conjunto de espejos de este tipo es que sólo quedaría enfocada la luz paralela al eje de la ordenación; la luz procedente de otras direcciones quedaría reflejada, saliendo del ojo. En consecuencia, lo que se necesita es una estructura que se comporte ópticamente con una pila de espejos en una configuración de borde de platillo, pero que pueda formar imágenes a partir de rayos luminosos que incidan en todas las direcciones.



ORDENACION EN CUADRADOS de las cajitas de espejos en el ojo compuesto de tipo reflector de la langosta *Palinurus vulgaris*, visualizada en esta micrografía. Los espejos de las cajitas cuadradas actúan como reflectores en esquina, los cuales tienen la propiedad de comportarse como un espejo único que se encontrara siempre perpendicular al rayo incidente. (Micrografía tomada por David Scharf.)

La solución al problema se halla en la estructura del propio ojo compuesto reflector. Este se halla cubierto por un conjunto de reflectores en esquina: son pares de espejos situados formando ángulos rectos. Estos reflectores tienen la propiedad de comportarse como un espejo sencillo que estuviera orientado siempre perpendicularmente al rayo incidente. (Los reflectores en esquina se utilizan en la navegación como reflectores de radar en las balizas flotantes: siempre devuelven el rayo del radar directamente al barco que lo transmite.) Por tanto, una ordenación de reflectores en esquina se comportará como la de bordes de platillo, pero lo hará cualquiera que sea la dirección de la que procede la luz. La única limitación importante de esta ordenación es que las cajas plateadas que componen el conjunto deben ser lo suficientemente profundas para que la mayoría de los rayos se encuentren con dos paredes, pero no tan profundas como para que la luz se refleje más de dos veces.

Experimentos realizados como modelos sugieren que las cajas deberían presentar una profundidad doble de su anchura, lo cual se aproxima muchísimo a la configuración real observada en los ojos de los langostinos y los cangrejos de río. Los espejos en disposición hexagonal no actuarían como reflectores en esquina y, consecuentemente, serían incapaces de hacer converger la luz. El empaquetamiento en forma de cuadrados de las facetas de los ojos de los crustáceos macruros no puede considerarse en absoluto como una característica secundaria del grupo; es crucial para el funcionamiento del sistema óptico.

Una conclusión lateral interesante que se desprende de estas observaciones es que la estructura del ojo compuesto, tanto si posee facetas hexagonales o cuadradas, puede darnos una pista importante en cuanto a la historia de la evolución de los crustáceos. El paso del ojo de superposición de tipo refractor al de tipo reflector exige un número tan elevado de modificaciones coordinadas que parece probable que ambos tipos sólo posean una relación muy lejana. Es de interés apuntar aquí que dos de los grupos de crustáceos más parecidos a los langostinos, los misidáceos y los eufausiáceos (estos últimos también llamados "krill", que es el alimento principal de las ballenas verdaderas), tienen ojos de superposición con facetas hexagonales. No es mi fuerte la taxonomía, pero creo no equivocarme al decir que este hecho sugiere que ambas familias de crustáceos se hallan, desde el punto de vista filogenético, mucho más alejadas de los macruros de lo que se entiende actualmente en la trayectoria que se supone para la evolución de los crustáceos. Por otra parte, los galateidos (las galateas, es decir, un grupo intermedio entre los macruros y los cangrejos verdaderos) tienen ojos de superposición de tipo reflector con facetas cuadradas, rasgo que los coloca filogenéticamente más cerca de los macruros.

Hace quince años, la mayoría de los zoólogos habrían aceptado la conclusión de que ya eran conocidos todos los mecanismos ópticos fundamentales utilizados por los organismos vivos y de que se conocían hacía medio siglo o más tiempo incluso. Ahora han surgido dos mecanismos nuevos, diferentes de los demás. Estoy bastante convencido de que hemos llegado al final del camino, pero, ¿quién sabe? Resultaría una sorpresa agradable comprobar que estoy equivocado.

El templo arrasado de Akhenaton

El faraón que deseaba que Egipto adorase un solo dios erigió imponentes monumentos en Karnak, que arrasaron sus sucesores. Lo que tales edificaciones eran se está recomponiendo ahora a partir de fragmentos

Donald B. Redford

En 1375 a.C. un insigne faraón terminó su reinado sobre Egipto. Era un miembro de la dinastía que hacía el número 18 entre las que habían detentado el poder real en el valle del Nilo desde que el régimen monárquico se había impuesto allí alrededor del 3000 a.C. Amenofis III había controlado el imperio durante casi cuarenta años de prosperidad. Su autoridad se hacía sentir mucho más allá de Egipto, desde el Sudán, en el sur, hasta Palestina, Siria y la frontera de Mesopotamia, en el este. Hubiera sido muy difícil predecir que al cabo de cinco años de la muerte del faraón, su hijo y sucesor Amenofis IV iba a revolucionar Egipto y que después, antes de terminar el siglo, casi todas las señales de la revolución real habían de quedar borradas.

El escenario principal de estos dramáticos acontecimientos fue la sede del poder real: Tebas, capital del reino situada a más de 500 kilómetros al sur del delta del Nilo. La capital constaba de dos centros de culto religioso en la ribera oriental del Nilo, Karnak y Luxor, y de una extensa ciudad de los muertos en la ribera occidental. Desde 1966 las investigaciones arqueológicas de la ribera oriental han contribuido a iluminar tanto los actos revolucionarios de Amenofis IV (que pronto cambió su nombre por el de Akhenaton con que hoy le conocemos) y las ásperas represalias de sus sucesores. Aquí voy a exponer lo que se ha averiguado en las últimas campañas de trabajo con particular referencia a la búsqueda de rastros del templo y de los otros edificios construidos por Akhenaton en Karnak.

En los años de formación de la religión egipcia los dioses eran divinidades locales. Osiris recibía culto en Busiris, Ptah en Menfis, la capital del Imperio Antiguo, y así sucesivamente. La divinidad local de Tebas era Amón-ra, representado por una figura humana con la cabeza y los cuernos de un carnero. La construcción del templo principal del dios en Karnak parece haber comenzado alrededor del año 2000 a.C. en tiempos del Imperio Medio. Una divinidad local semejante, establecida en On (la Heliópolis de los griegos), era Atón-ra. Aunque venerado en forma de toro, como Ptah lo era en Menfis, Atón-ra estaba considerado como divinidad solar. A comienzos de la época del Imperio Medio (2445 a.C.) Atón-ra, o simplemente Ra, el dios sol, era ya identificado con muchas deidades locales egipcias, y en particular, en Tebas, con Amón-ra.

La mayor parte del templo de Amón en Karnak fue construida a comienzos del Imperio Nuevo (a partir del 1580 a.C.), y sobre todo en los años que siguieron a la expulsión por Amosis, fundador de la Dinastía XVIII, de los intrusos hicsos, que habían gobernado Egipto durante las dinastías XV y XVI. De los diez grandes pilonos, o puertas ceremoniales de piedra, que tenía el templo, seis, desde el tercero al octavo, fueron construidos durante la Dinastía XVIII (1580-1350 a.C.). Durante su largo y próspero reinado, Amenofis III (1411-1375) hizo también sus adiciones a los esplendores del templo del dios del sol. Tres pilonos más, el segundo, noveno y décimo, fueron construidos en diferentes circunstancias durante la Dinastía XIX (1350-1200 a.C.).

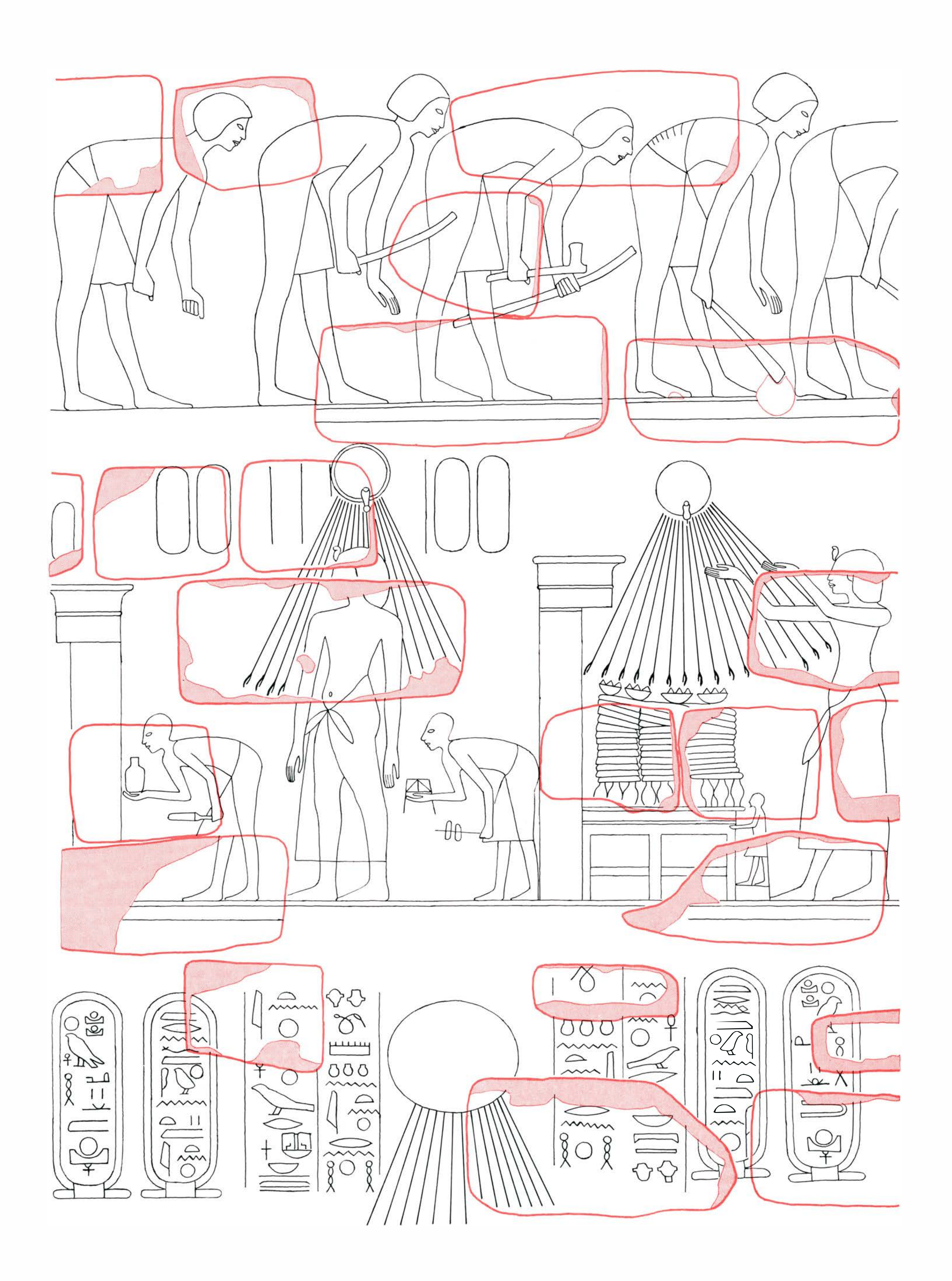
Amenofis III, "Rey del Alto y Bajo Egipto, Nebmaatra, Hijo del Sol" (para darle su título oficial), era un hombre de gusto refinado para grandes edificios, antigüedades y mujeres. Casó con varias esposas de alcurnia, parte de ellas por lo menos por razones de estado. Una era una princesa de Babilonia, otra una princesa de Arzawa (en lo que es ahora el suroeste de Turquía) y una tercera una princesa de Mitanni (moderno Kurdistán).

Su reina principal, sin embargo, era una mujer del pueblo llamada Teye. Ostentaba el significativo título de "Gran Heredera", y conservó el rango de "Esposa del Gran Rey" hasta el final del reinado de Amenofis. Teye sobrevivió a su marido por lo menos un decenio. De esta unión nacieron seis hijos, cuatro hembras y dos varones. Esta rama principal de la familia amenófida proporcionó a Egipto dos generaciones de gobernantes.

El menor de los dos hijos de Teye, llamado también Amenofis, era débil y deforme. Algunos sabios modernos, basando su diagnóstico en representaciones nada halagadoras de su físico, creen que el príncipe sufría de una anormalidad de pituitaria, conocida hoy en día como síndrome de Froehlich. A la muerte inesperada de su hermano mayor, este joven sin pretensiones se encontró convertido en príncipe heredero. Cuando su padre falleció, víctima de varias dolencias, en 1375 a.C., el príncipe, casi desconocido de todos excepto de su madre, ascendió al trono como Amenofis IV.

Si, como su padre antes que él, el nuevo rey iba a ser "Hijo del Sol", fue ésta una responsabilidad que él no tomó a la ligera. Al principio de su reinado, decidió adorar a un solo dios: el Disco Solar, Re-horakhte. Durante su primer año de reinado, las representaciones animales de la divinidad quedaron prohibidas: las cabezas de halcón y cabezas de

CULTO AL DISCO SOLAR, representado dos veces en la franja intermedia de esta reconstrucción de un relieve de tres frisos, que muestra a Akhenaton acompañado de dos sacerdotes, bañado en los rayos del Disco Solar (*izquierda*) y solo, con los brazos levantados (*derecha*), cubriendo un altar de ofrendas de alimentos para el Disco Solar. En la hilera superior, cortesanos haciendo reverencias, provistos de las varas de su cargo, presentan sus respetos a Akhenaton. Los jeroglíficos de la franja inferior identifican a la divinidad soiar con su manifestación el el Disco Solar. Las líneas coloreadas muestran el perfil de cada uno de los sillares de arenisca; el color claro señala la erosión. (Bunji Tagawa).



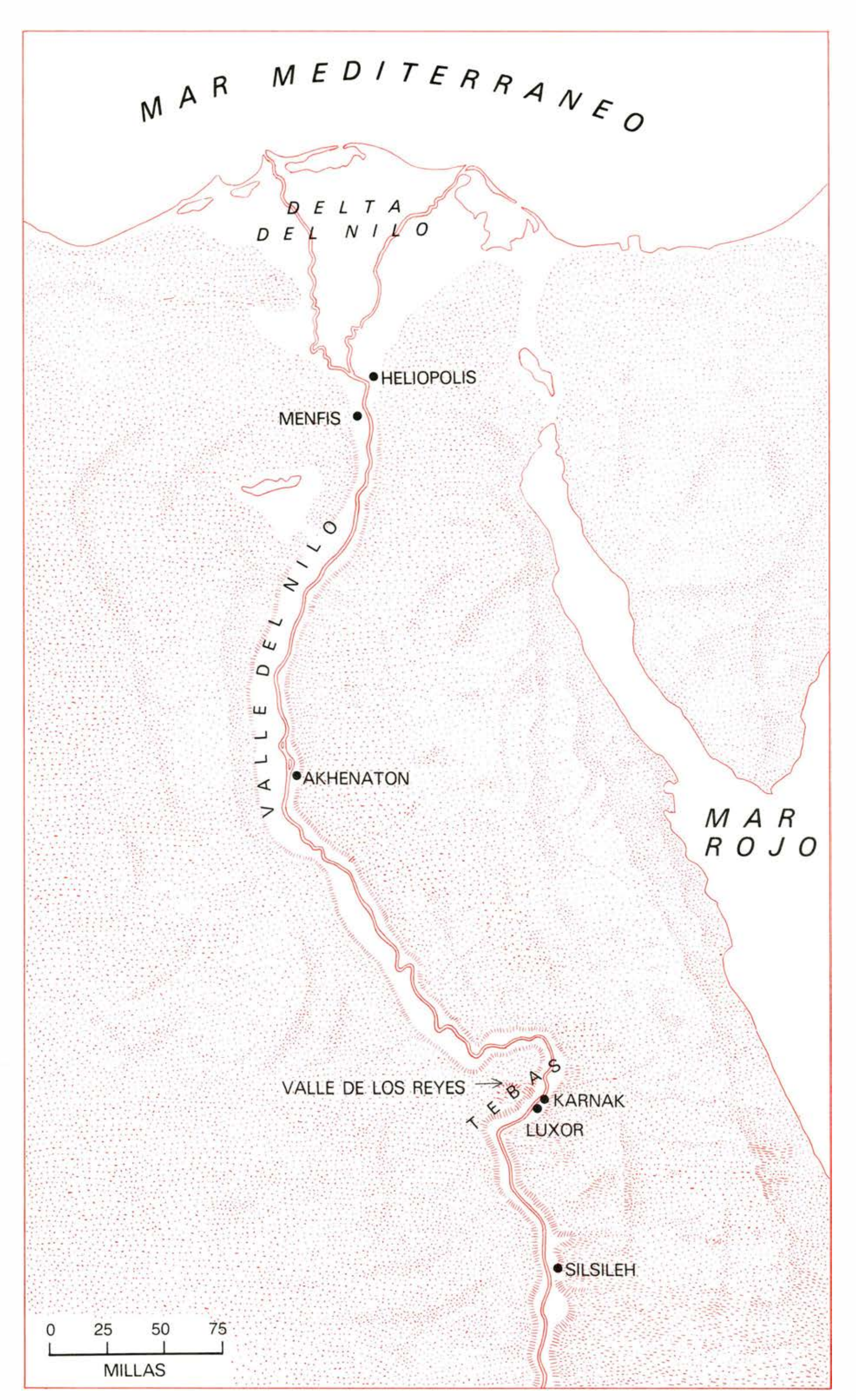
carnero superpuestas a cuerpos humanos dejaron de ser aceptables como representaciones del sol divino. En un acto paralelo de administración, el nuevo faraón decretó que las rentas de todos los demás dioses egipcios fuesen abonadas a un solo tesoro, el del Disco Solar. Los reales decretos lastimaron más que a nadie a los sacerdotes de Amón-ra de Tebas, que habían sido los mayores beneficiarios financieros de más de un siglo de conquistas exteriores.

El nuevo régimen del faraón se expresó en obras tanto como en decretos. En el segundo y tercer año de su reinado, Amenofis IV celebró un magno jubileo en la capital regia. Por lo menos ocho edificios de cantería fueron erigidos o comenzados para la celebración, y el lugar elegido por el nuevo señor para la erección de la mayor parte de los edificios nuevos fue el templo de Amón-ra en Karnak.

Conocemos los nombres de cinco de estos edificios. El más complicado era uno llamado Gen-po-aton, "(el) Disco Solar es Hallado". Grabadas en relieve en las paredes de arenisca del edificio había escenas de la fiesta del jubileo y escenas que mostraban a la esposa del rey, la reina Nefertiti, y una o dos de las hijas que por entonces había tenido, haciendo ofrendas al sol. Otro edificio, posiblemente construido dentro del primero, era el "Hat-ben-ben", "la Mansión de la Piedra Ben-ben". El título es antiguo; durante mucho tiempo había sido aplicado a un templo muy primitivo de Heliópolis, la ciudad donde la divinidad solar Atón-ra era venerada. Los relieves de las paredes de arenisca de la "Mansión" mostraban a Nefertiti en el acto de hacer ofrendas al sol; aquí no está acompañada por sus hijas. La "Mansión" no se contruyó hasta el cuarto año del reinado del nuevo monarca, por lo menos.

Dos santuarios se construyeron poco después del jubileo. Uno se llamaba "Exaltados son por siempre los Monumentos del Disco Solar". Los relieves esculpidos en las paredes de arenisca del templo "Exaltado" representaban muchas escenas de la vida doméstica de palacio; las del santuario "Robusto" desplegaban filas de braseros para ofrendas al sol, procesiones de carrozas, soldados y escenas de palaciegos. Unos cuantos grabados dejan ver al rey haciendo ofrendas al sol. El nombre de un quinto edificio "Caseta del Disco Solar" es también conocido por inscripciones. Nada se sabe de los nombres, decoraciones o situación de los otros tres edificios regios levantados por Amenofis IV en los primeros cinco años de reinado.

Fue en su quinto año de trono cuando el nuevo monarca cambió su nombre por el de Akhenaton ("Util al Disco Solar"). Ordenó el cierre del templo de Amón-ra en Karnak y dispuso, como toque adicional, que el nombre de Amón-ra fuese raspado de todos los monumentos templarios. En compañía de la reina y de las hijas de ambos (ahora tres) Akhenaton se mudó de Tebas a una nueva ciudad que había construido en la orilla oriental del Nilo, a unos 225 kilómetros al norte. La nueva ciudad (la moderna Amarna) recibió el nombre de Akhetaton, "Horizonte del Disco Solar".



EL VALLE DEL NILO AL NORTE DE ASUAN está encajonado por el desierto al este y al oeste. La capital regia, Tebas, constaba de dos centros ceremoniales en la ribera oriental del Nilo, Karnak y Luxor, y de una dilatada ciudad de los muertos en la ribera occidental. Silsileh, al sur, es el emplazamiento de la cantera de arenisca que proporcionó los sillares empleados en la construcción de los edificios ceremoniales de Akhenaton en Karnak. Akhetaton, al norte, es la capital construida cuando Tebas fue abandonada. Más al norte se hallan Heliópolis, donde Atón-ra era objeto de culto, y Menfis.

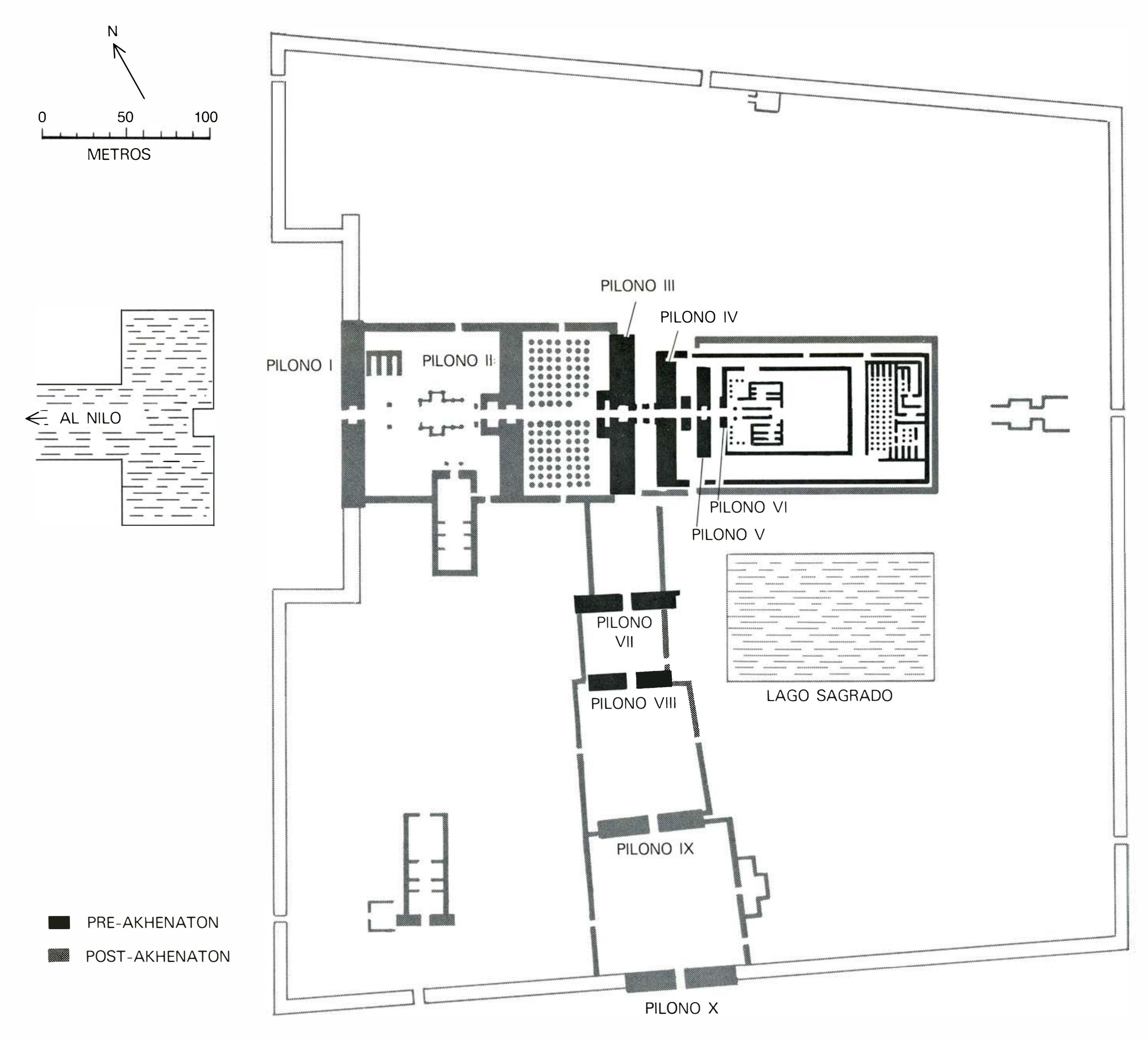
MUROS DE LAS CASAS de un barrio residencial ubicado al este del templo de Amón-ra, desenterrados en 1976 y 1977. El barrio fue ocupado en tres fases sucesivas a lo largo de un período de unos 500 años de duración en el primer milenio a.C. Vivían allí jóvenes sacerdotes y escribas.

PILARES RECTANGULARES, desmantelados casi hasta el nivel fundacionai. Marcan la columnata meridional del templo del jubileo de Akhenaton. La orientación es hacia el oeste. El muro sur del templo, tras la columnata, fue demolido en su totalidad, si bien quedaron muchos escombros.

Allí celebró su duodécimo aniversario de reinado con un gran despliegue de tributos y una reunión de dignatarios extranjeros.

Akhenaton no fue un digno émulo de su padre como diplomático. Los hititas minaron la frontera siria de Egipto e incluso conquistaron al aliado de Egipto, Mitanni, sin ninguna respuesta enérgica por parte de Akhenaton. En 1358 a.C. durante el año decimoséptimo de su reinado, el reformador religioso murió.

Durante los once años siguientes, el trono fue ocupado sucesivamente por los yernos de Akhenaton, Smenkhkare (1360-1357) y el famoso Tutankhamon (1357-1347). Smenkhkare parece haber apoyado las reformas internas de su suegro durante su breve reinado, pero después de tres años en el trono, Tutankhamon fue cediendo, hasta el punto de trasladar la sede real a Tebas desde la nueva ciudad construida por Akhenaton. Tutankhamon fue sucedido en 347 a.C. por Aye, el último de los Amenófidas. A la muerte de Aye, en 1344 a.C., el jefe del ejército de Tutankhamon, Horemheb, asumió la autoridad real como fundador de la XIX Dinastía. Durante un breve período, Horemheb, que reinó desde 1344 hasta 1315 a.C., respetó el compromiso amenofídico con el Disco Solar, e incluso pudo haber hecho adiciones a los edificios de Akhenaton en Karnak. No pasó mucho tiempo, sin embargo, antes de que la situación cambiase. De un modo concienzudo, que casi sugería animosidad personal, los templos del sol del palacio real en Akhetaton (la ciudad), fueron desmantelados sistemáticamente y arrancados sus cimientos. El mismo destino corrieron los edificios de Akhenaton en Tebas y en Luxor, y, en particular, la serie de construcciones vecinas al templo de Amón-ra en Karnak. El culto al sol de Akhenaton dejó de existir; su nombre y los de sus tres sucesores inmediatos fueron incluso borrados de la lista oficial de los reyes de Egipto.



TEMPLO DE AMON-RA EN KARNAK. La serie de construcciones importantes consignadas en este plano incluye aquellas emprendidas después de la demolición de los monumentos de Akhenaton en honor del Disco Solar, situados en las proximidades y que emplearon la cantería de los muros derribados. El núcleo de la puerta ceremonial rotulada Pílono II en el plano, y los cimientos de la sala columnada situada al este del mismo, se componen casi enteramente de cantería del templo del jubileo de Akhenaton que había sido demolido, cantería recogida y vuelta a emplear por los cons-

La obra de demolición no constituyó un despilfarro. Por ejemplo, cuando Horemheb añadió los pílonos segundo, noveno y décimo al restablecido templo de Amón-ra en Karnak, los materiales de relleno, que proporcionaron la base y núcleo del pílono segundo, y la cimentación de una gran sala columnada, justamente al este del pílono, consistieron casi exclusivamente en los sillares de arenisca del templo del jubileo de Akhenaton "(el) Disco del Sol es Hallado". La base del pílono noveno consistió, en su mayor parte, en sillería de arenisca rescatada de los dos santuarios demolidos, el "Exaltado" y el "Robusto".

Cuando Ramsés II (1304-1239 a.C.), tercer faraón después de Horemheb, emprendió construcciones en las proximidades de Tebas, sus canteros trasladaron a Luxor otros varios miles de bloques de arenisca que habían sobrado de la demolición de los edificios de Akhenaton en Karnak. Más de un milenio después, cuando Roma gobernaba Egipto, cierta cantidad de bloques sobrantes de Akhenaton fueron a parar a la construcción privada; algunos aún pueden verse en los muros de antiguos edificios del Cairo, a unos 450 kilómetros de Karnak. Más triste aún de contar, muchos de los bloques esculpidos con mayor elegancia fueron a parar a las manos de coleccionistas particulares, tanto de Egipto como del extranjero.

En el día de hoy se conservan en las cercanías de Tebas más de 40.000 bloques de Akhenaton: hay siete almacenes llenos de ellos en Karnak y otro almacén en Luxor. La mayoría, si no todos ellos, fueron transportados a las construcciones de Akhenaton desde las canteras de arenisca de Silsileh, en la ribera oriental del Nilo, a 125 kilómetros al sur de Karnak. Fueron labrados a su tamaño en cantera por miles; el encargo general para las unidades era de bloques de 22 centímetros de espesor, 25 centímetros de alto y 52 centímetros de largo. Las paredes de los edificios de Akhenaton en Karnak fueron construidas poniendo los bloques en hiladas, alternando a soga y tizón, y sirviéndose de yeso como mortero. Una vez que la pared estaba terminada, canteros especializados realizaban la decoración grabada en relieve y otras inscripciones por la cara interior.

Tebas y su comarca, que comprende el dramático Valle de los Reyes, donde la escarpadura del Sáhara orla la ribera occidental del Nilo, ha sido escenario de una activa investigación arqueológica desde mucho antes del descubrimiento de la tumba de Tutankhamon por Lord Carnavon y Howard Carter en 1922. Misiones francesas y alemanas trabajaron en la zona durante el siglo XIX y lo mismo antes que después de la Primera Guerra Mundial. En realidad fue la obra de restauración de los franceses en el templo de Amón-ra en Karnak, en las décadas de 1920 y 1930, la que por vez primera puso de manifiesto la reutilización de los bloques de Akhenaton a millares en la base y núcleo del segundo pílono y de la sala columnada.

El hecho de que la mayoría de los bloques mostrasen parte de una escena relivaria en una de las caras demostraba que no eran simples cascotes, sino restos de lo que habían sido edificios de elevado rango. Merced al llamativo estilo artístico introducido en época de Akhenaton y tan fácil de reconocer, se impuso pronto la convicción de que la gran mayoria de la arenisca tallada era una parte, si no la casi totalidad, de los edificios que el reformador había erigido en Karnak en los primeros cinco años de su reinado. Al ir apareciendo más y más bloques, la simple tarea de almacenarlos para ponerlos a salvo de robos, para no hablar de la tarea de analizarlos, fue haciéndose cada vez más formidable.

Casi nada se hizo con respecto a este problema hasta veinte años después de terminada la Segunda Guerra Mundial, cuando R. W. Smith, funcionario retirado del servicio extranjero americano y asociado de investigación del Museo Universitario de la Universidad de Pennsylvania, puso en marcha el proyecto del templo de Akhenaton. El proyecto estaba patrocinado por el Museo Universitario a base de fondos del programa de financiación exterior de la Smithsonian Institution. El enfoque de Smith era a un tiempo sencillo y heroico: fotografiar a una misma escala cada bloque grabado, reseñar la cara grabada de modo apropiado para un análisis por computador, y hacer luego que un programa de computación se hiciese cargo del enorme número de pruebas necesarias para ajustar cada cara grabada a todas sus vecinas. Una vez hecho esto, la reunión de las fotografías sería una tarea relativamente sencilla.

En los once años transcurridos entre este enorme esfuerzo analítico y su terminación en la primavera del año pasado, Smith y su grupo internacional de expertos han logrado resultados muy superiores a los sueños de cualquiera. Treinta y seis mil bloques han sido fotografiados y miles de escenas reconstruidas. Los resultados permiten seguir los pasos al progreso de Akhenaton en el desarrollo de, por lo menos, los aderezos exteriores de su monoteísmo radical. También arrojan luz sobre la vida diaria durante los cinco años en que permaneció en Tebas, e incluso suscitan interesantes problemas acerca de sus relaciones con su esposa la reina.

El edificio más antiguo de Akhenaton construido en Karnak, probablemente durante el primer año de su estancia en el trono, fue una puerta monumental, un pílono del que no quedan restos. Sus bloques de arenisca eran mucho mayores que los del tamaño establecido para la cantería de muros. Entre las zonas labradas que han sobrevivido hay una larga inscripción. En este texto, el nuevo

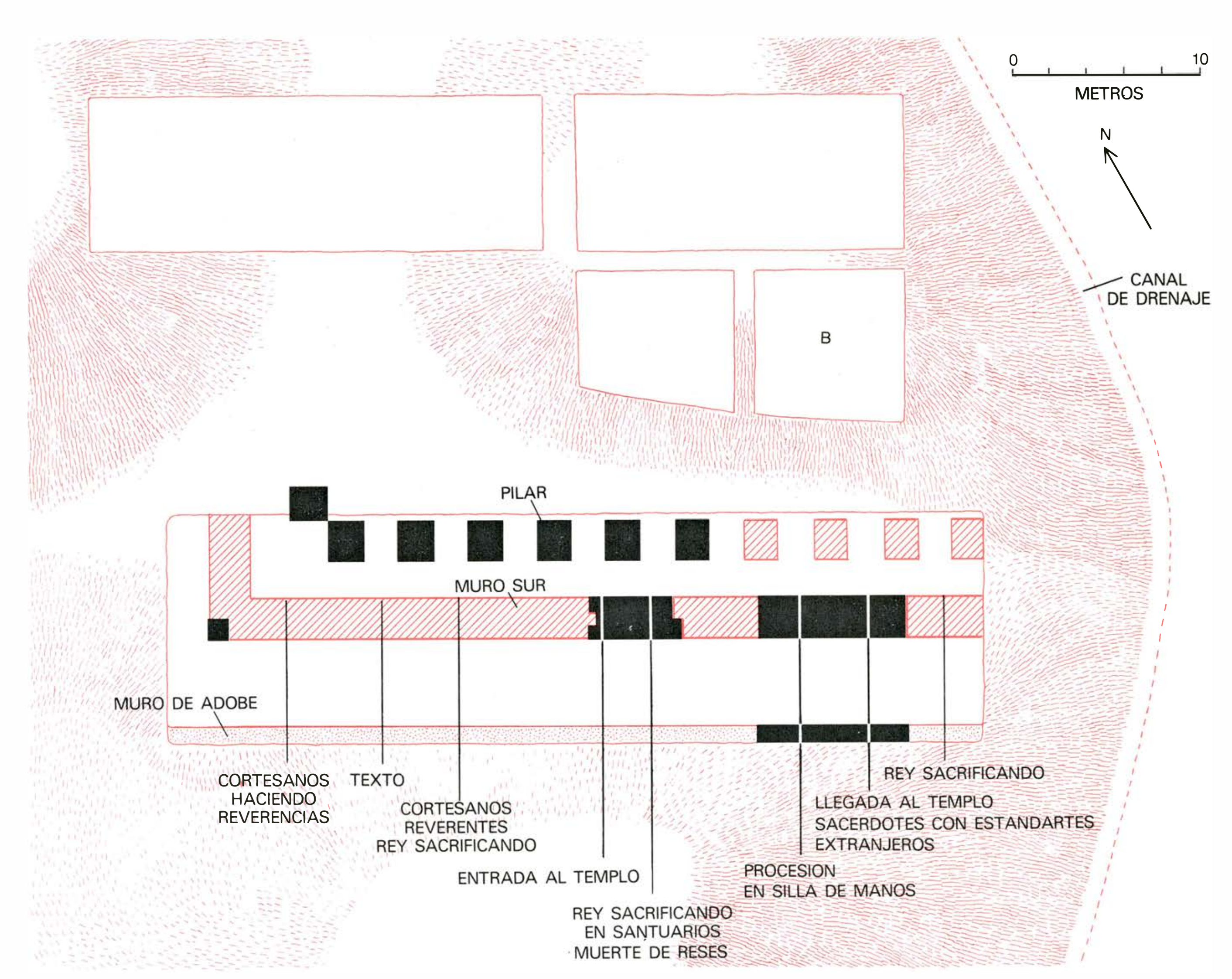
tructores del faraón Horemheb. El permiso de excavación dado al Proyecto del Templo de Akhenaton en 1974 (*color*) se extiende un kilómetro al norte y al sur (*véase la ilustración siguiente*).

rey pondera las cualidades incomparables del dios solar a expensas de todos los demás cultos divinos de Egipto. La mayor parte de las representaciones del dios solar que aparecían en el pílono seguían la tradición iconográfica más temprana de representar al dios con un cuerpo humano y una cabeza animal. Sin embargo, antes de que concluyese la decoración del pílono Akhenaton había introducido su nueva imagen del sol: un simple disco que emite muchos rayos delgados, cada uno de ellos terminado en forma de una mano humana. Por tanto, el abandono de la representación semiantropomorfa del dios solar puede fecharse con probabilidad antes del fin del primer año de reinado de Akhenaton como nuevo faraón.

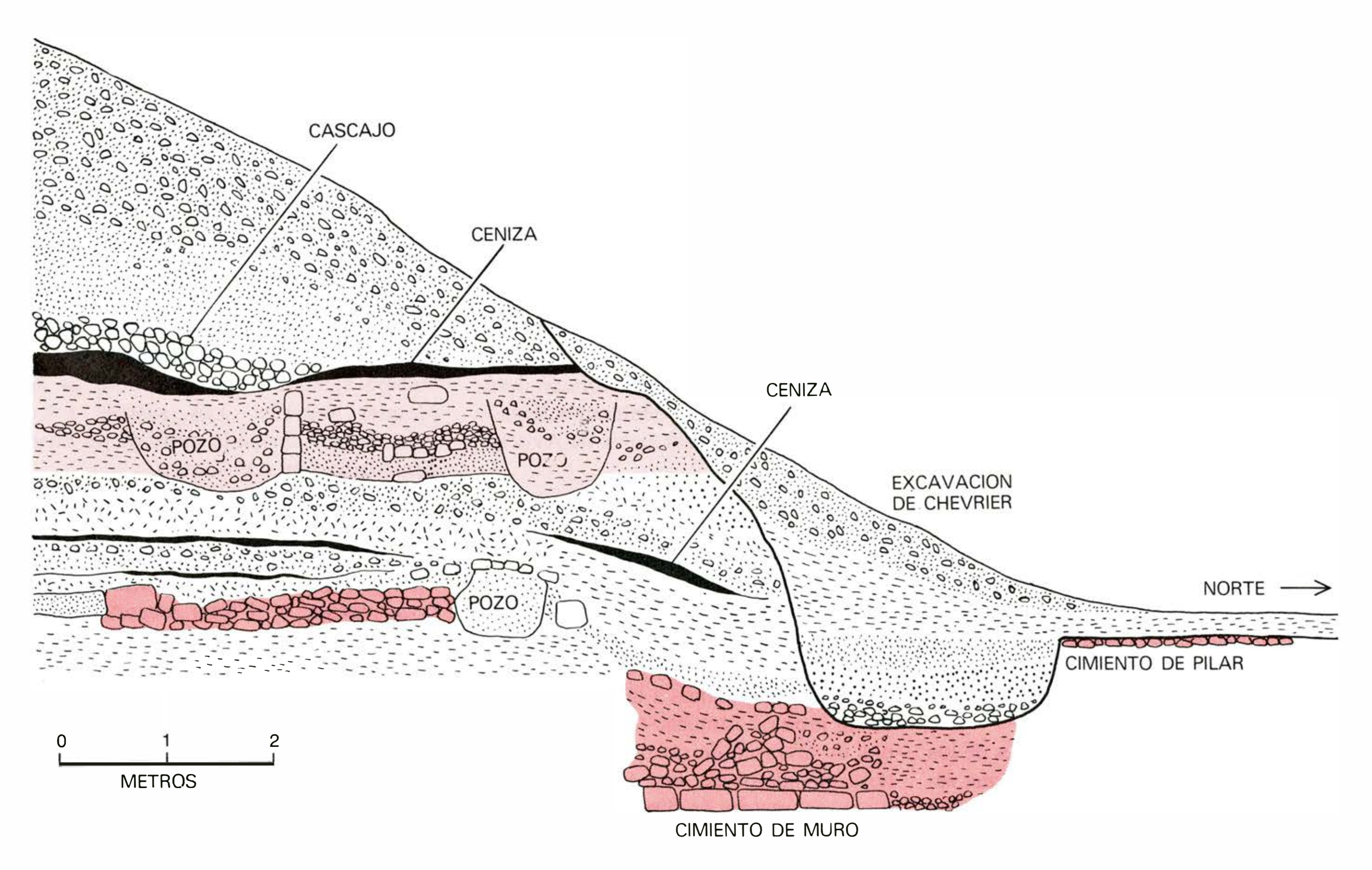
En cuanto a las relaciones entre Akhenaton y Nefertiti, en los relieves del templo del jubileo, y también en las de los dos santuarios construidos poco después, el rey está representado mientras hace ofrendas al Disco Solar, unas veces solo, otras en compañía de la reina, que está figurada a escala reducida. Cuando no antes del cuarto año de su residencia en Tebas se añadieron una columnata al templo del jubileo y la Mansión de la Piedra Ben-ben, las escenas de ofrenda habían pasado a ser casi de la reina. En los relieves de la columnata se ve a Nefertiti con una o con sus dos hijas mayores, pero, por lo regular, con la mayor de ellas, Meretaton. En los muros de la Mansión, es la reina la que hace ella sola las ofrendas.

No se ve claro por qué esto tuvo que ser así. Quizás estas imágenes de la reina como mediadora con el cielo tengan relación con otra singularidad de Karnak: la evidente ocupación de la reina en otras actividades asociadas normalmente con las prerrogativas masculinas de un faraón. Nefertiti no sólo está representada haciendo sacrificios, sino también recibiendo a prisioneros en cadenas e incluso aporreando las cabezas de víctimas estilizadas.

Al cumplir su décimo año, el proyecto de investigación del templo de Akhenaton había dado unos resultados sorprendentes, pero desequilibrados. Se sabía mucho de la decoración relivaria, de su estilo y de sus temas, pero



ZONA DE LA CONCESION inmediatamente al este del templo de Amón-ra, excavada parcialmente por investigadores anteriores. El trabajo iniciado en 1975 incluyó la exploración de la parcela de ocho metros cuadrados señalada con una *B* en la parte superior derecha de este plano y una segunda parcela del mismo tamaño al sur (no dibujada). La continuación del trabajo en 1976 amplió la excavación en la línea *B* de parcelas y finalmente reveló los primeros indicios del templo del jubileo (*abajo, a la derecha*): los cimientos de la pared sur del templo y los escombros de su demolición, indicios del muro de cinta exterior, de adobe, que rodeaba el templo y la base de uno de los pilares que sostenían el techo. Las excavaciones siguientes han descubierto una serie de bases de pilares; éstos formaron antaño una columnata por el interior de la pared meridional del templo. Los rectángulos negros señalan las bases de los pilares que hemos localizado; las zonas negras de los muros están documentadas de modo similar por restos de cimientos y detritus de la cantería rota del muro mismo. Trozos de escenas conservados en fragmentos de cantería, que corresponden a escenas recompuestas merced al estudio de la sillería conservada del templo, aparecieron a lo largo del muro en el orden indicado. (Dibujo de Bunji Tagawa).



LAS EXCAVACIONES INICIALES en la concesión de Karnak–Este tuvieron que bajar cortando unos 2500 metros cúbicos de vaciaderos de anteriores excavadores, antes de alcanzar los restos de un barrio residencial (*color claro*) que fue detruido dos veces entre el siglo VII y el III a.C. En agosto de 1976 los obreros llegaron a una zona de arenisca fragmentada, y más allá de esa zona, a las bases de pilares rectangulares (*color intenso*).

nada se había averiguado de los templos y de otros edificios cuyas paredes habían desplegado los relieves, como tampoco sobre su arquitectura o incluso sobre sus plantas. Quienes habían puesto interés en extirpar todas las pruebas de las obras de Akhenaton habían desmantelado sus edificios tan a conciencia, que ni una huella de los mismos podía verse por ninguna parte en Karnak. Los especialistas no tenían una idea clara, ni aún en fecha tan tardía como 1974, de dónde se había levantado ninguno de los edificios del reformador; incluso discutían, a menudo acaloradamente, qué funciones podían haber ejercido los edificios. Fue, por tanto, una suerte que en el otoño de 1974 el proyecto obtuviese un permiso de la Organización Egipcia de Antigüedades para efectuar excavaciones a cien metros en Karnak.

La financiación de la nueva empresa corrió a cargo del Smithsonian y, más tarde, del Programa Killam del Consejo del Canadá. En abril de 1975, un grupo bajo mi dirección comenzó las excavaciones a 100 m. al este del gran muro de cerramiento que delimita el templo de Amón-ra en Karnak. Elegimos la zona porque en 1925 el Ayuntamiento de Luxor había abierto allí una zanja de drenaje. Durante aquella obra los excavadores habían puesto al descubierto una hilera de estatuas de Akhenaton colosales que estaban enterradas, así como restos de un muro construido con los bloques de arenisca que caracterizan a los edificios de Akhenaton. A lo largo de los años siguientes, el inspector de Antigüedades de Egipto, Henri Chevrier, trató de seguir la hilera de colosos hacia el este y hacia el oeste, sin resultados dignos de mención. Su tarea no era fácil, porque durante más de un siglo aquella zona se había utilizado como vaciadero por casi todos los arqueólogos que habían trabajado en Karnak.

Nuestra concesión, de 250 metros de ancho por casi un kilómetro de longitud, abarca la vieja excavación de Chevrier y el vaciadero. Como ésta es la única zona en que hay constancia de haber existido algo construido por Akhenaton, era un punto de partida lógico, de modo que al comenzar la campaña de 1975 delimitamos dos parcelas cuadradas, de ocho metros de lado, una al sur y otra al norte, en las faldas de la antigua excavación de Chevrier.

La excavación con maquinaria está aún por introducir en la arqueología egipcia, de modo que no tuvimos más remedio que recurrir a los métodos tradicionales de palear y llevar la tierra en espuertas. Pese a ello, en las tres primeras campañas logramos sacar unos 5000 metros cúbicos de tierra de nuestros cuadrados (que fueron ampliados en la campaña de 1976 mediante tres cuadrados más al sur, y uno al norte). ¡Más de la mitad de esta cantidad consistía en los vaciados de otros excavadores!

En todas las zonas excavadas por nosotros encontramos, por debajo del nivel de la zanja de 1925, paredes de bastante entidad, de casas que datan aproximadamente desde el siglo séptimo al tercero a.C. Pronto se vio con claridad que durante este período nuestro sector había formado parte de un barrio residencial. Las pocas inscripciones halladas sugerían que la mayor parte del barrio había pertenecido a la clase media y probablemente habia estado habitado por inspectores, sacerdotes jóvenes y escribas.

Los ceramistas también habían estado ocupando una zona, como lo indicaban los restos de dos alfares. El análisis de los fragmentos de restos de animales (recuperados por flotación) indica que la nutrición de los residentes era

rica en carne de vaca; más del setenta y cinco por ciento de los huesos de animales son de vacuno.

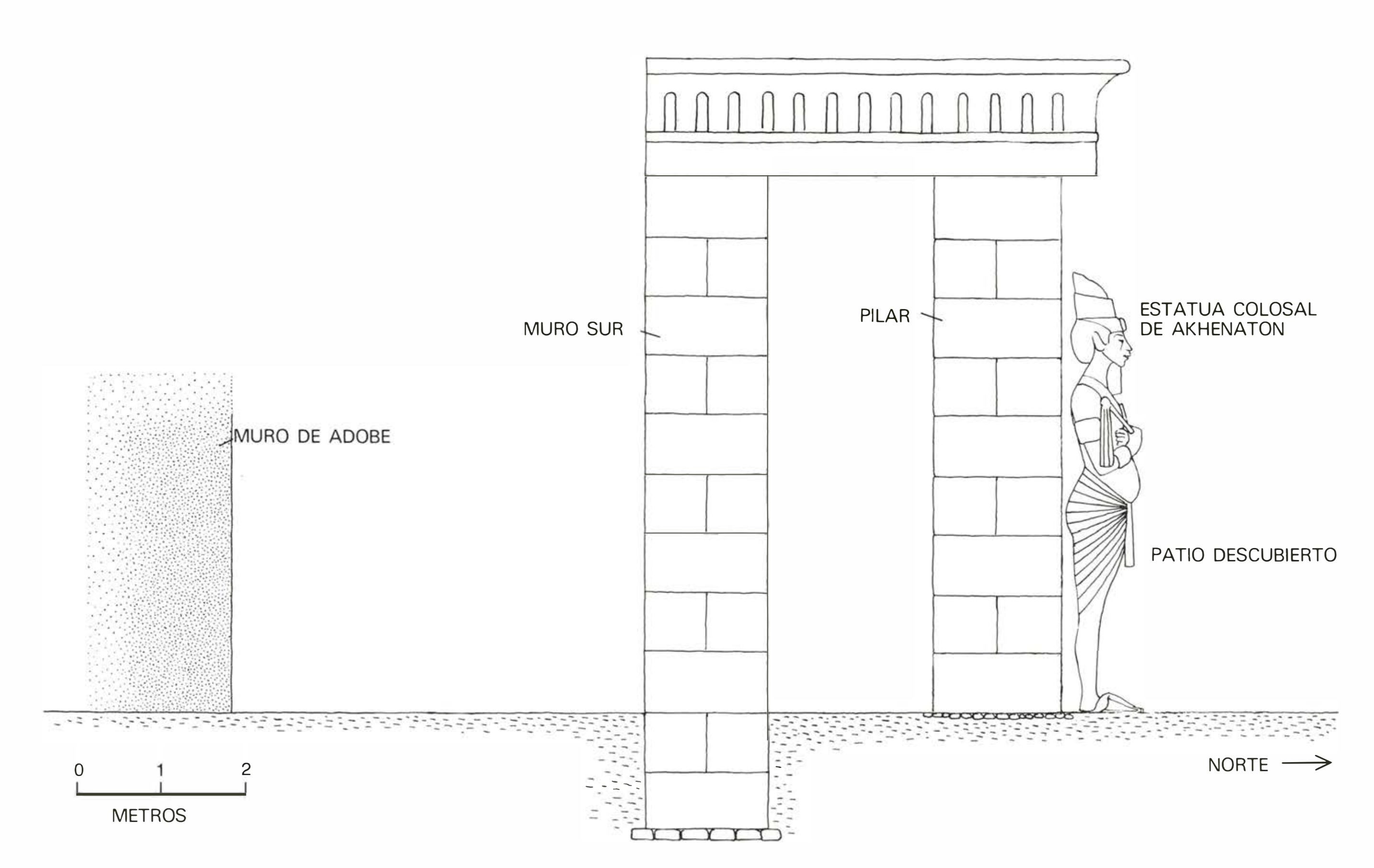
Durante los 500 años de ocupación, el poblado había sido destruido dos veces. La prueba de estas perturbaciones estaba clara en el nivel de construcción más antiguo, y era imposible no percibirla en el nivel medio. Aquí los suelos estaban sembrados de escombros: vasos cerámicos hechos añicos y otras señales de actividad violenta. En contraste con estos episodios anteriores, el nivel superior de construcción, último en ser ocupado, fue abandonado en unas condiciones que sugerían que sus habitantes por fin habían tenido tiempo de hacer las maletas. Los pisos estaban perfectamente barridos y pocos utensilios domésticos habían quedado atrás. La datación precisa de cualquiera de estos niveles de ocupación es difícil, pero se hace tentador atribuir la violenta destrucción del nivel medio a la invasión persa del 525 a.C.

En nuestra búsqueda de señales de los edificios de Akhenaton, tuvimos que seguir cavando a través de los restos que databan del primer milenio a.C. El primer metro por debajo del más antiguo de los tres niveles de ocupación estaba colmado de relleno, debido en su mayor parte a un basurero antiguo. Los tiestos mezclados con este material podían atribuirse a los primeros tres o cuatro siglos del milenio. En un sector descubrimos los restos de un edificio hecho de ladrillo cocido. No había inscripciones que ayudasen a identificar el edificio, pero algunos ladrillos llevaban la estampilla de un sumo sacerdote de Amón-ra llamado Menkheperre, que se sabe floreció desde 1020 hasta 1000 a.C. Habíamos llegado a una distancia de tres siglos de Akhenaton.

Hasta agosto de nuestra segunda campaña no logramos pruebas de la venganza que Haremheb y sus sucesores habían tomado sobre Akhenaton. Descendiendo en nuestra excavación, con calores de verano de 48,9 grados centígrados, llegamos a la parte superior de un montón inmenso de bloques partidos de arenisca de un muro, esparcidos desordenadamente en un abanico de unos seis metros en dirección este-oeste. Lo siguiente en descubrirse, situados paralelamente en relación con los bloques esparcidos del muro, fueron las bases de varios grandes pilares rectangulares que medían dos metros por 1,80.

Tumbado cara arriba en el montón de arenisca había un fragmento de sillar que llevaba parte de un relieve labrado; sus líneas resultaban muy familiares. No era para menos. El fragmento formaba parte de una escena de silla gestatoria que habíamos conocido desde hacía tiempo merced a nuestros nueve años de trabajo ajustando las fotografías de los bloques murales. Los sillares rotos, pues, habían formado parte de un muro construido por Akhenaton.

¿Pero qué muro? Las dos campañas siguientes proporcionaron la respuesta. En 1976 recobramos cerca de cien fragmentos más de relieves de una zanja de doce metros, y el total ascendió a 250 cuando la zanja fue ampliada otros treinta metros el año pasado. La mayor parte de los relieves de los sillares pudo ser identificada al momento como procedente del más complejo de los edificios de Akhenaton en Karnak, el templo Gem-po-aton. Para redondear la cuestión, sacamos posteriormente dos ins-



PARED SUR del templo del jubileo de Akhenaton. Hallábase un tiempo inmediatamente detrás de una hilera de pilares cuadrados, a unos cinco metros por el interior de un muro de cerramiento de adobe (*izquierda*). Una estatua colosal de Akhenaton, situada delante de cada pilar, miraba hacia el patio abierto y con piso de tierra apisonada. Una estatua de este género (ninguna de ellas se ha encontrado en la excavación) aparece en el dibujo a línea de este perfil. El perfil mantiene la orientación sur-norte de la sección transversal de la excavación. La altura de la pared de arenisca y de los pilares no es conocida, de modo que el techo que aparece aquí es supuesto, pero debe haber sido más alto que las estatuas de seis metros que cubría.

cripciones que incluían el nombre del templo.

Al término de nuestra campaña de 1977 se podía contemplar la configuración general de lo que habíamos descubierto en Karnak. Habíamos llegado a la pared sur y a la columnata interior de una gran patio descubierto, con un piso de tierra apisonada. Los pilares de la columnata estaban asentados a trechos de dos metros; apoyada en cada pilar y mirando al norte hacia dentro del patio, se alzaba antaño una estatua colosal de Akhenaton. Estas eran las figuras que se habían descubierto por vez primera en 1925, al hacer la zanja de la alcantarilla, y que fueron después recogidas a finales de la década de 1920 y durante la de 1930.

El muro de arenisca decorado con relieves se había alzado a unos dos metros al sur de la hilera de pilares; su zanja de cimentación bajaba hasta metro y medio por debajo del piso de tierra apisonada. Los relieves habían sido labrados en la cara interna del muro, donde quedarían ocultos parcialmente por los pilares de la columnata. Las escenas esculpidas constituían una secuencia continua de decoración que se extendía desde el suelo hasta la techumbre. Cinco metros al sur del muro de sillares de arenisca, y siguiendo una línea paralela al mismo, se alzaba una segunda pared hecha de adobes; esta pared debe haber formado parte de un recinto sagrado que rodeaba el templo del Disco Solar.

Nuestro trabajo nos ha permitido reconstruir con cierto grado de certeza el modo preciso como el templo fue destruido. Primero, se derribaron las estatuas sobre sus caras. Después, los pilares de la columnata fueron desmontados hasta casi el nivel del suelo. El muro de circunvalación de adobe fue arrasado antes de emprender el ataque del muro de arenisca, evidentemente desde el sur. Se desmontó el muro de arenisca sillar por sillar, haciendo saltar el mortero a golpes de martillo. Un espeso montón de detritus de arenisca se fue acumulando a lo largo del muro, y en él, y alrededor, cayeron muchos sillares despedazados. El desmantelamiento de la pared, a diferencia del derribo de los pilares, se prolongó hasta más abajo del nivel del suelo.

Hay buenas razones para creer que el abandono de los fragmentos de sillares de arenisca coincide con los puntos en donde éstos cayeron. Esta hipótesis de trabajo nos ha proporcionado una ocasión única: la posibilidad de reconstruir la secuencia de las escenas que decoraban los muros del templo. Justamente esta información sobre la secuencia de las escenas, a diferencia de su composición interna, era lo que nos faltaba en nuestros mosaicos fotográficos.



FRAGMENTO DE UN BLOQUE recuperado de entre las ruinas del muro del templo del jubileo. Muestra parte de la escena que decoraba la cara del muro: dos portadores del trono de Akhenaton.

Por consiguiente, hemos cartografiado con cuidado la posición de cada uno de los doscientos cincuenta y tantos fragmentos de arenisca decorada que hemos recuperado hasta ahora. En la mitad de los fragmentos, aproximadamente, la decoración está lo bastante clara, aunque a veces es muy pequeña, para determinar qué clase de escena y en ocasiones a qué escena precisa corresponde el fragmento. Por ejemplo, hemos podido sacar la conclusión de que todos los fragmentos que descubrimos al principio pertenecían al tema que más se repite en el templo: una procesión de la familia real dirigiéndose al templo desde su palacio. Los personajes principales se mueven de derecha a izquierda y Akhenaton lleva la corona blanca que simboliza al Alto Egipto, esto es, al valle del Nilo en cuanto éste se opone a las tierras septentrionales del delta del Bajo Egipto, que a su vez estaban simbolizadas por una corona roja. La corona representativa del Egipto meridional era naturalmente la decoración idónea para la decoración de la pared meridional del templo.

Nuestros fragmentos se relacionan con escenas que representan etapas en el acercamiento al templo; la carrera está cubierta de dignatarios extranjeros y de sacerdotes portadores de estandartes. Los fragmentos de la zanja situada más hacia el este corresponden a una escena en la que se ve el muro de cinta de adobe; Akhenaton está dentro del recinto haciendo una ofrenda al Disco Solar. De modo semejante, la más occidental de las zanjas ha suministrado fragmentos pertenecientes a una escena de la procesión cuando ésta sale del palacio real.

Esta secuencia de escenas, desde el palacio al interior mismo del recinto sagrado, ocupaba un tramo de pared de unos 13 metros de largo. Lo que pudo haber habido más allá de la última escena que conocemos por el este, viene indicado por lo que hemos podido reconstruir más hacia el oeste. En la esquina donde la pared meridional y la occidental se encuentran, están las figuras de hombres haciendo reverencias. Más allá de los mismos, hacia el este, hay un texto que por desgracia está demasiado mal conservado para poder leerlo. Seguidamente, se encuentra el pílono del templo y, a continuación, una escena en dos niveles: sacrificio de toros abajo y ofrendas rituales encima.

Si esa es la última escena de la secuencia, no habría por qué sorprenderse

de encontrarla donde está: el artista se limitó a repetir el motivo de la procesión. Quizás era su intención transmitir el constante movimiento y el colorido de varias procesiones consecutivas al templo. Después de todo, el jubileo de Akhenaton, el acontecimiento que está representado, duraba muchos días y comportaba muchas visitas al templo.

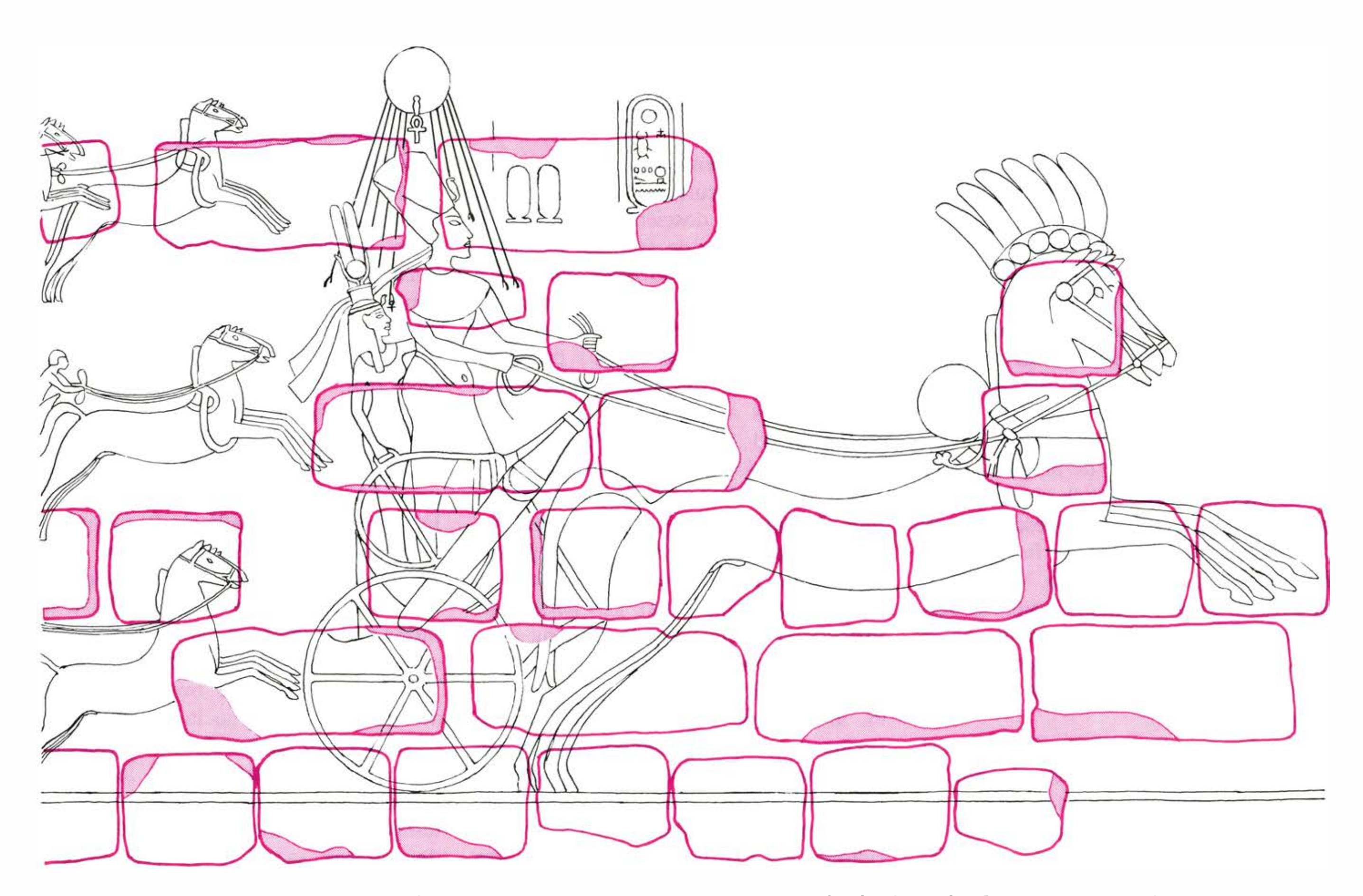
Conforme las excavaciones prosiguen caba esperar que hagan más excavaciones sobre el destino de los edificios de Akhenaton. Nuestra concesión de un kilómetro debería, sin embargo, suministrar información de un alcance más amplio. Por ejemplo, la investigación del establecimiento urbano de finales del primer milenio a.C. podría ocuparnos durante los años venideros. No es frecuente que uno tenga ocasión de excavar un sector grande de una ciudad regia. ¿Cómo era aquí el panorama ecológico? ¿Cuál era la base económica de la ciudad? ¿Cómo se hallaba planificada? Estos y otros muchos interrogantes vienen con facilidad a la mente.

Ya durante la campaña de 1978 hemos comenzado a vislumbrar respuestas a muchas de nuestras preguntas. Por ejemplo, hemos descubierto que la columnata y la pared decorada del templo se extendían a lo largo del lado occidental del patio del templo. También aquí las escenas representan procesiones del rey y de la reina en su camino del palacio al templo. También se ve a los sacerdotes y cortesanos haciendo libaciones.

Hay algunos indicios, conforme adelantamos hacia el norte siguiendo el flanco occidental del templo, de que podemos estar acercándonos a un lugar donde antaño se alzaba un pilono. La anchura exacta del patio del templo en sentido norte-sur todavía se nos escapa, aunque nuestro magnetómetro parece haber captado señales del muro norte en un punto situado a 90 metros del lado meridional del patio.

Cualesquiera que sean las enseñanzas que Karnak proporcione, siempre tendrá especial importancia cualquier nueva información que suministre sobre Akhenaton. El faraón reformista es quizás el más conocido y posiblemente el más atrayente de todos los gobernantes antiguos. Es probable que esto se deba a que se le contempla sobre todo como el primer defensor del monoteísmo. Por su veneración de un solo dios se ha hecho de Akhenaton un precursor de Moisés, de Cristo y de Mahoma. En rigor, el contenido espiritual y ético de la religión hebrea y del cristianismo primitivo faltan por completo en el programa de Akhenaton, lo mismo que los aspectos solares de su religión son tan ajenos a las enseñanzas de Moisés como de Cristo. La iconoclasia de Akhenaton tuvo distantes ecos en el nacimiento del Islam, pero el más próximo aliado espiritual de su culto solar fue el efímero culto al "Sol Invictus" del emperador Aureliano en el siglo III d.C.

El patrocinio por parte del faraón de un estilo artístico sorprendentemente moderno no lesiona su reputación; la verdad es que parece uno de nosotros. Sea o no correcta nuestra valoración del hombre, no cabe duda de que sus realizaciones como reformador, actuando desde arriba, son únicas en épocas primitivas. Hay ironía en su fama, pues las acciones de Akhenaton poco más produjeron que una onda en la corriente de la larga historia antigua de Egipto.



FARAON Y REINA, Akhenaton y Nefertiti (detrás de él, a escala reducida) en un carro tirado por dos caballos dirigiéndose desde su palacio a uno de sus templos dedicados al Disco Solar en Karnak. Los rayos del sol descienden sobre la pareja real. Los otros tres carruajes están ocupados por cortesanos y oficiales de escolta. Las inscripciones indican que esta escena decoraba una de las paredes del templo de jubileo de Karnak. Futuras excavaciones en la concesión de Karnak podrán determinar su ubicación anterior. El más próximo aliado espiritual de Akhenaton fue Aureliano.

Nebulosas gaseosas

Esas espectaculares nubes dentro de nuestra galaxia constituyen sedes de formación estelar. La radioespectroscopía indica, entre otras cosas, que las nubes pertenecen a sistemas, mucho mayores, de gas y polvo

Eric J. Chaisson

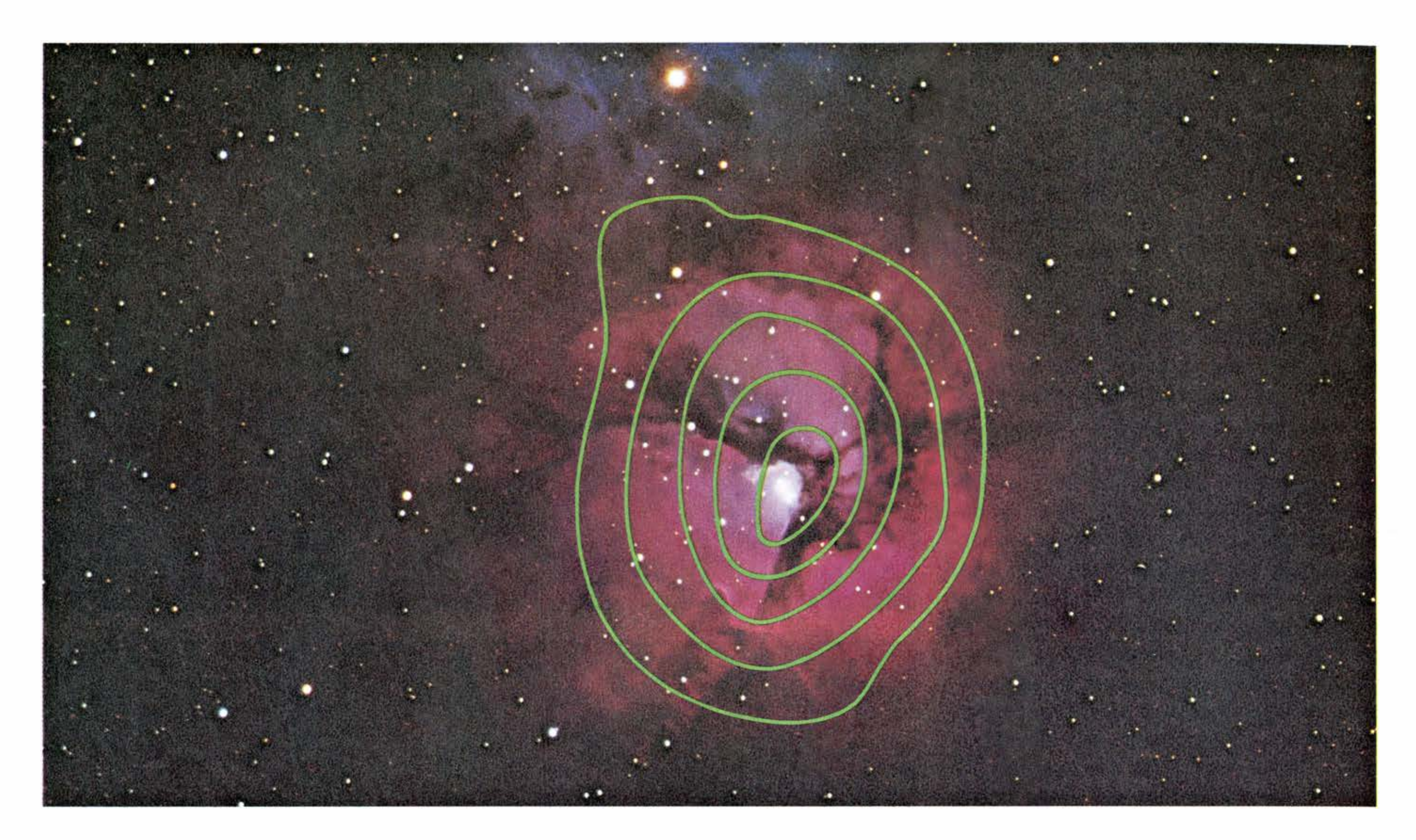
Los astrofísicos suelen estar de acuerdo sobre los grandes rasgos de la teoría de la formación de estrellas. Resumiendo, una estrella nace de la manera siguiente. Al principio, una nube difusa y fría de gas y polvo llena una considerable región del espacio interestelar. Dentro de la nube, algunos átomos quedan ligados, al principio por accidente. Supuesto que su fuerza gravitatoria combinada sea mayor que cualquiera de las fuerzas que se oponen, comienzan a atraer otros átomos de la proximidad. Con el tiempo, la acción gravitatoria continuada forma una densa concentración de materia y procede a comprimirla todavía más. El calor debido al rozamiento hace aumentar la temperatura del gas comprimido por la acción de la gravedad. Tras algunos millones de años, el gas alcanza una temperatura crítica de unos diez millones de grados Kelvin, punto en el cual los núcleos de hidrógeno comienzan a fundirse en la parte interna de la naciente estrella. La liberación de enormes cantidades de energía termonuclear procedente de la "combustión" del hidrógeno detiene el colapso gravitatorio de la estrella, que ulteriormente radia energía desde su superficie hacia el espacio en forma de ondas electromagnéticas, incluyendo las de luz visible.

Pese a la casi universal aceptación del cuadro anterior, hasta hace poco disponía de muy escasa base observacional. Hace algunos años, cierto número de observatorios astronómicos emprendieron programas destinados a examinar regiones del espacio de las que se creía podían conducir a la formación de estrellas. La atención se dirigió inicialmente hacia las nebulosas gaseosas y, más tarde, hacia su entorno interestelar. Esas partes, visualmente espectaculares, de la galaxia son ricas en gas y polvo, y se sabe que contienen estrellas jóvenes en comparación (algunas de no más de un millón de años). Sin embargo, los lugares donde coexisten estrellas tan jóvenes y protoestrellas todavía más primitivas tienden con frecuencia a ser polvorientas y, por tanto, obscuras; están parcial o totalmente ocultas a la vista de astrónomos que trabajan con telescopios ópticos a causa de la eficaz difusión de las ondas luminosas por el polvo interpuesto. Por fortuna, las ondas de radio penetran tales detritus interestelares mucho más fácilmente, haciéndolas superiores a las ondas luminosas para explorar el interior tanto de nebulosas gaseosas parcialmente obscurecidas como de nubes galácticas poco luminosas y totalmente obscurecidas. Pasaré revista aquí a algunas interesantes observaciones radioastronómicas hechas en los últimos años sobre una selección representativa de nebulosas gaseosas y sus entornos estelares. Tales investigaciones podrían muy bien suministrar la clave de una comprensión más completa del proceso de la formación de estrellas.

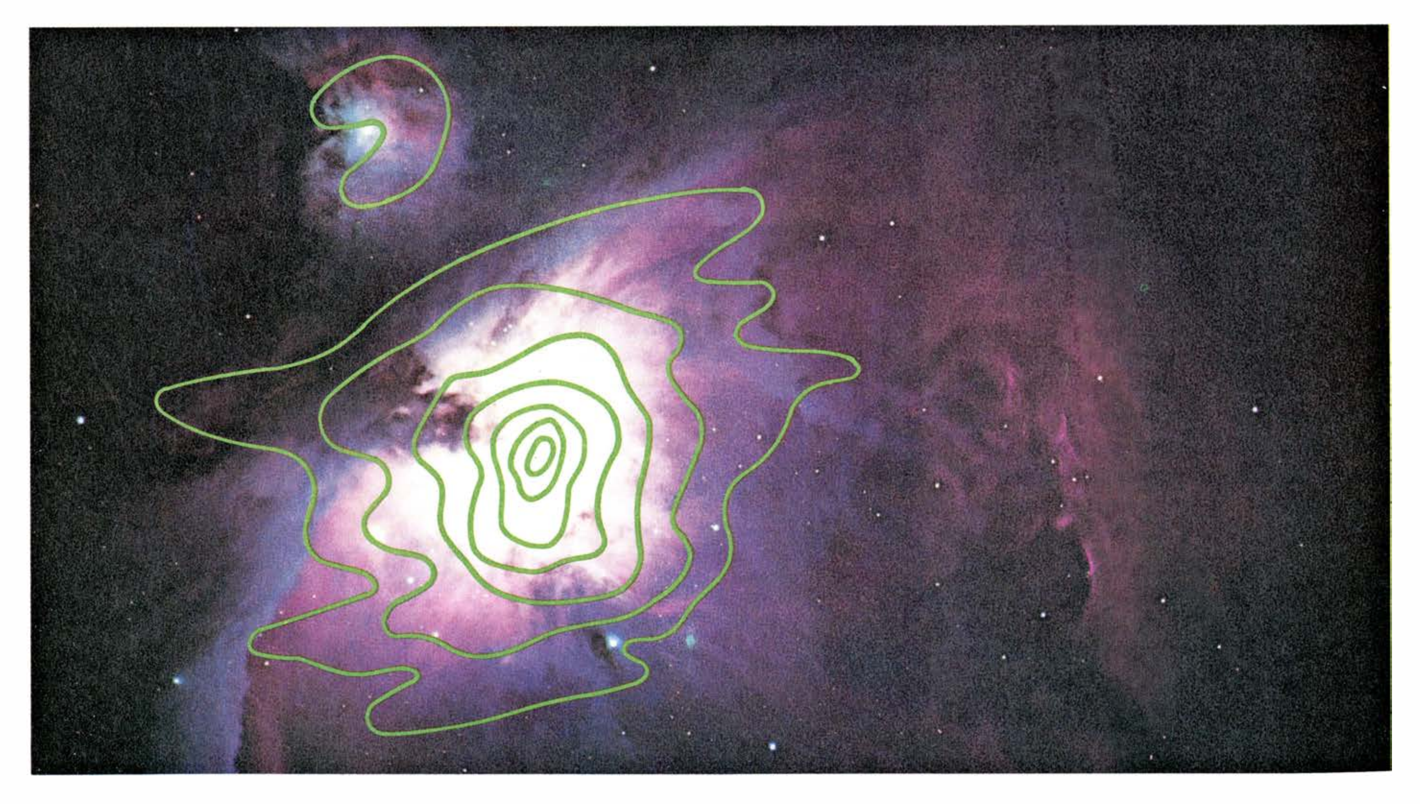
Una nebulosa gaseosa es un volumen más o menos esférico de plasma, o gas ionizado, que rodea a una estrella muy caliente o a un cúmulo de estrellas, cada una de las cuales tiene una temperatura efectiva en su superficie de más de 30.000 grados Kelvin. Tales estrellas blanco-azuladas son mucho más calientes que nuestra propia estrella, el Sol, cuya temperatura en superficie es aproximadamente de sólo 6000 grados. Las estrellas calientes del primer tipo emiten copiosas cantidades de radiación ultravioleta, la cual es capaz de interaccionar con la materia circundante. De todos los fotones de la radiación electromagnética emitida por tal estrella, sólo los fotones ultravioleta, de frecuencia comparativamente alta, tienen más energía de la que se necesita para mantener ligado un electrón a su núcleo. Como resultado, la absorción de radiación ultravioleta por átomos neutros, o no ionizados, causa el que ciertos electrones sean arrancados de sus núcleos, creando un plasma casi perfecto alrededor de la estrella o estrellas centrales. La energía liberada por la interacción de un fotón ultravioleta y un átomo se cede al electrón liberado en forma de energía cinética, con lo que éste queda libre para vagar en el interior del gas ionizado. El comportamiento de tales electrones no ligados, llamados fotoelectrones, es crítico por lo que respecta a la observación de las nebulosas. Pueden manifestar su presencia de tres maneras diferentes.

En primer lugar, tras varios cientos de años de errar, un electrón libre puede con el tiempo acercarse lo suficiente a un núcleo (generalmente un núcleo de hidrógeno, o protón aislado) para ser recapturado, formando de nuevo un átomo neutro. Al comienzo del proceso de recombinación, el electrón suele unirse al sistema atómico ligado en un estado energético superior o excitado, descendiendo a niveles de energía más bajos a un ritmo medio de aproximadamente un nivel por segundo, con lo que alcanza el estado fundamental, o nivel de más baja energía, bastante rápidamente. Durante las fases sucesivas de la recombinación, se emiten fotones a frecuencias específicas, características de la estructura del átomo. Las transiciones de bajo nivel (es decir, las que tienen lugar entre unos pocos de los primeros estados cuánticos principales del átomo) son responsables de las familiares líneas de emisión del hidrógeno y el helio que aparecen en los espectrogramas ópticos de objetos estelares. Las transiciones entre niveles de energía adyacentes más altos del mismo elemento dan lugar a ondas de radio más largas, que también pueden escapar de la nebulosa y ser analizadas espectroscópicamente.

En segundo lugar, los electrones no ligados chocan, a menudo, con átomos de elementos escasos, tales como oxígeno ionizado sencilla o doblemente, o azufre ionizado con una carga, los cua-



NEBULOSA TRIFIDA, designada M20 en el catálogo de Messier de objetos celestiales extensos, se representa aquí de dos maneras. La imagen fotográfica, hecha con el telescopio de cuatro metros del Observatorio Nacional de Kitt Peak, registra las diversas longitudes de onda de luz visible emitidas por la nebulosa como resultado de la recombinación de electrones libres con distintos átomos en el volumen de gas ionizado que rodea a la estrella central excitadora. Las isolíneas verdes, trazadas a partir de datos obtenidos con la radioantena de 37 metros del Observatorio Haystack, muestran la distribución de radioondas emitidas por la nebulosa a la frecuencia de ocho gigahertz. En este caso, las líneas de intensidad constante crecen uniformemente hacia dentro, hacia la posición conocida de la estrella central, coincidente con la forma de la nebulosa visible. (Messier fue quien primero catalogó las nebulosas.)



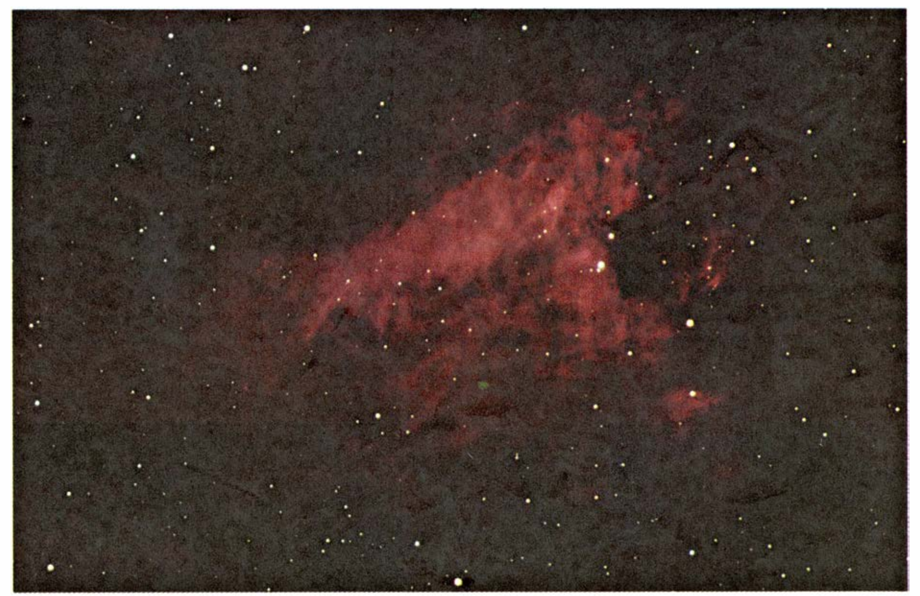
GRAN NEBULOSA DE ORION (M42), análogamente representada en ambas formas, óptica y de radio. En este caso, sin embargo, el mapa de radiointensidad muestra que partes importantes de la nebulosa quedan ocultas a la visión de los astrónomos que trabajan con telescopios ópticos. El radiomapa de M42, o Gran Nebulosa de Orión es más detallado que el de M20 porque, para esta observación, la antena de Haystack se hizo funcionar a la frecuencia de 24 gigahertz, y la frecuencia más elevada produjo mejor resolución angular. Casi ciertamente, están naciendo nuevas estrellas en o cerca de estas dos nebulosas gaseosas representativas. En todas las fotografías astronómicas que acompañan a este artículo, el norte está arriba y el este a la izquierda.

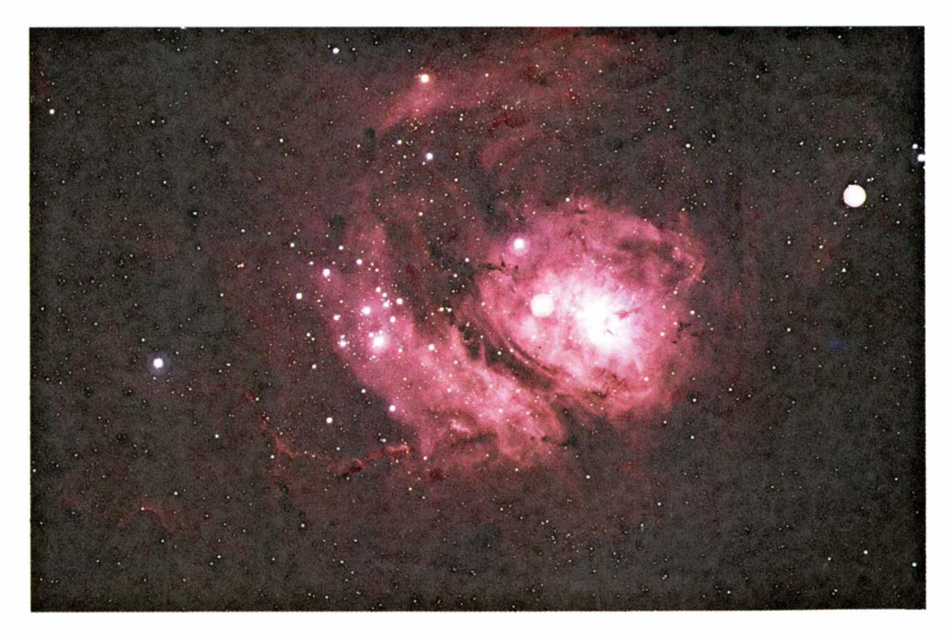
les tienen niveles de energía muy bajos. (Los iones, más abundantes, de hidrógeno y helio no tienen niveles de energía comparablemente bajos.) El proceso de choque puede ser suficientemente violento para transferir energía del electrón libre al ion, poblando así algunos estados iónicos justo por encima del estado energético fundamental. En principio, los choques pueden también despoblar los niveles excitados de un ion, pero, por lo general, los choques son bastante raros en el tenue ambiente interestelar, y como resultado, una vez que un ion ha sido excitado por choque a un nivel más alto, es mucho más probable que descienda radiativamente al nivel fundamental emitiendo un fotón de baja energía capaz de escapar de la nebulosa. Tales transiciones, que se llaman prohibidas porque sus líneas espectrales no se pueden detectar en los mucho más densos gases que se encuentran en los laboratorios terrestres, son en parte responsables de los visualmente espectaculares colores de las nebulosas gaseosas; el oxígeno simplemente ionizado emite líneas prohibidas en la parte violeta del espectro visible, el oxígeno doblemente ionizado en el verde, el azufre simplemente ionizado en el rojo, y así sucesivamente.

Por último, los choques inelásticos, o de cesión de energía, entre los propios electrones libres, pueden distribuir con rapidez (en unos minutos) la energía inicialmente ganada por el electrón en toda la región del gas ionizado, estableciendo un estado termodinámico llamado distribución de Maxwell, caracterizado por una temperatura media de la nebulosa de alrededor de 10.000 grados. (Esta temperatura del gas es en realidad una medida de la velocidad media de los electrones en la nebulosa.) De esta manera, la estructura térmica de la nebulosa se mantiene generalmente en un régimen estacionario de equilibrio entre el calentamiento por fotoionización y el enfriamiento por los procesos de recombinación y choque que generan radiación de baja energía capaz de escapar de la nebulosa. Es de importancia más que incidental el hecho de que sean los iones escasos (aquellos cuya abundancia es al menos 1000 veces menor que la del hidrógeno y el helio) los que sirven para controlar la nebulosa termostáticamente, manteniendo su temperatura generalmente entre 5000 y 15.000 grados.

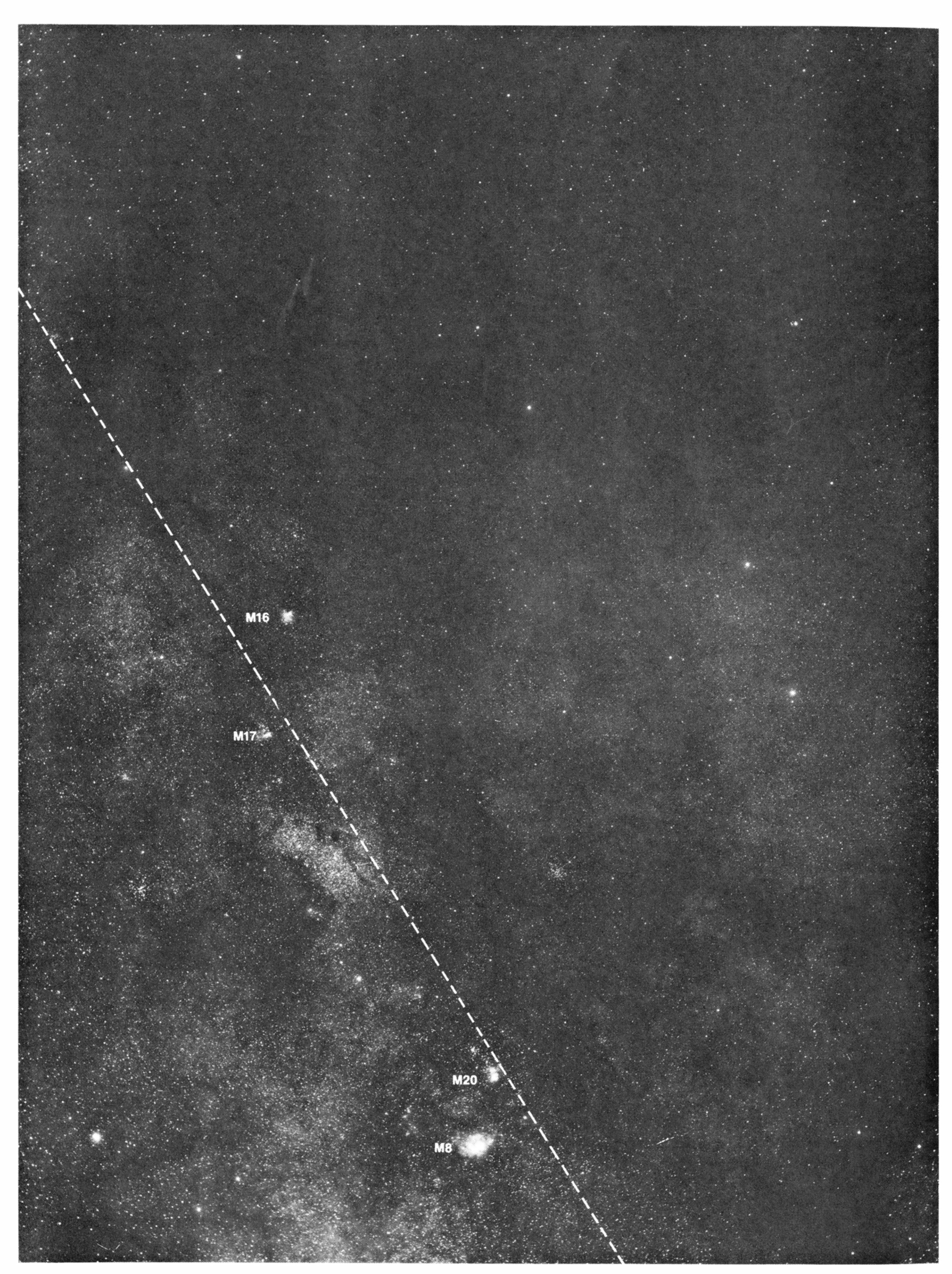
Además del equilibrio térmico que se establece entre los diversos mecanismos de calentamiento y enfriamiento, las nebulosas gaseosas se caracterizan por estar en equilibrio de ionización. Como el tiempo que suele tardar un electrón en recombinarse (algunos cientos de años)



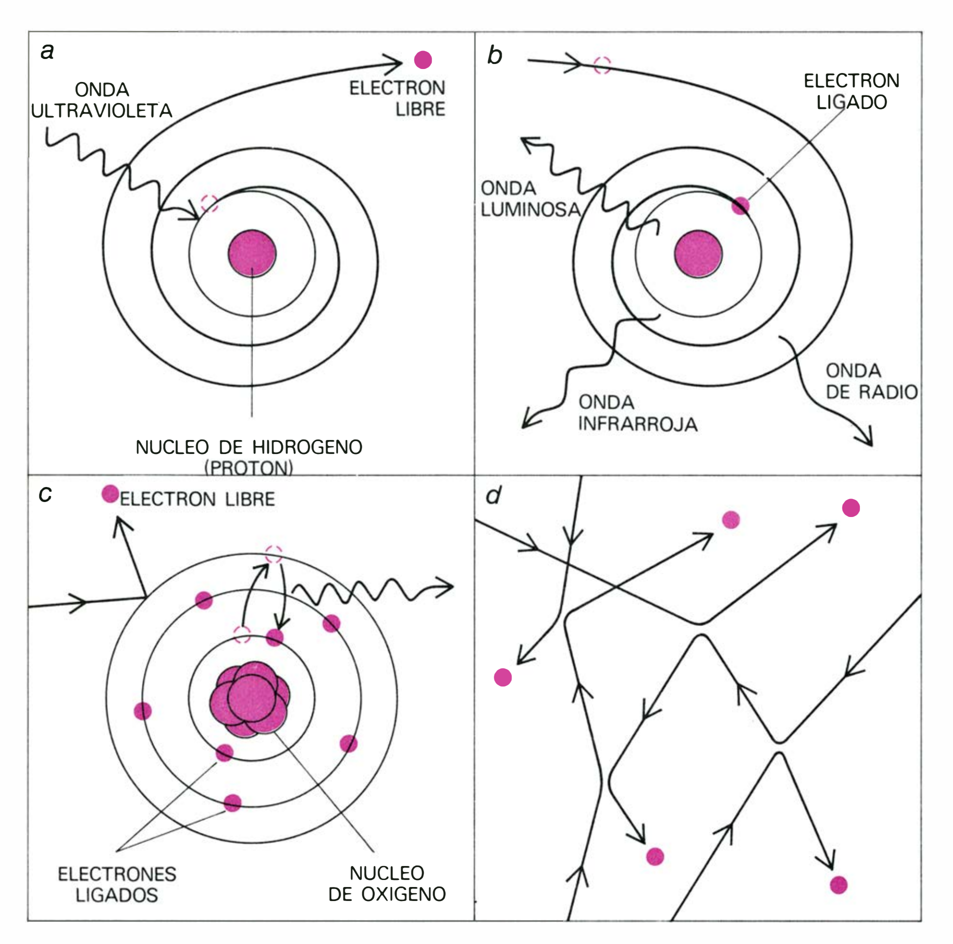




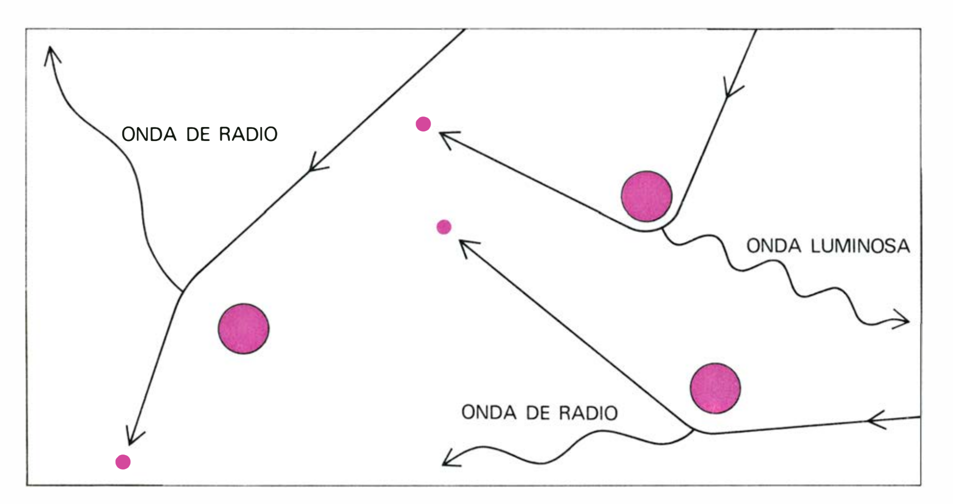
UNA REGION DE NUESTRA GALAXIA RICA EN NEBULOSAS, vista en la fotografía en blanco y negro y gran angular a la derecha, hecha con una cámara de vigilancia del cielo de 7,6 centímetros en el observatorio austral de la Universidad de Harvard, sito en la estación de Bloemfontein en Africa del Sur. La toma se ha hecho hacia el centro de la galaxia, al nordeste de la constelación de Sagitario; la línea blanca de



trazos indica el plano central de la galaxia. La fotografía se extiende a una sección angular del cielo extraordinariamente grande, que mide 26 grados en declinación (verticalmente) y 19 grados en ascensión recta (horizontalmente). Contiene varias nebulosas notables, incluyendo las tres que se muestran en color a la izquierda. De arriba abajo, son M16, la Nebulosa del Aguila; M17, la Nebulosa Omega; M8, la Nebulosa de la Laguna. (M20, también indicada en la toma con gran angular, se reproduce en color en otra parte de este artículo.) Las fotos en color se hicieron con el telescopio de cuatro metros de Kitt Peak.



LOS ELECTRONES LIBRES pueden interaccionar con otras partículas en el gas nebular ionizado de tres maneras diferentes, una vez que se han liberado individualmente del único protón de los núcleos de átomos de hidrógeno por la acción de la radiación ultravioleta de una estrella central (*a*). Primero, un electrón libre puede volver a ser capturado por un protón, formando un nuevo átomo de hidrógeno (*b*); a medida que el electrón salta hasta el nivel de energía más bajo del átomo, parte de la energía cinética del electrón se emite en forma de radiación electromagnética a longitudes de onda características (primero a baja frecuencia en forma de ondas de radio, después a frecuencia intermedia en las longitudes de onda del infrarrojo, y, finalmente, a longitudes de onda ópticas de alta frecuencia). Segundo, un electrón libre puede comunicar parte de su energía cinética por choque con ciertos átomos, tales como el oxígeno ionizado con una sola carga (*c*); el choque excita a un electrón dentro del ion a un estado de mayor energía, desde el cual puede regresar a su estado normal de energía, emitiendo energía en forma de luz visible. Tercero, los electrones libres pueden interaccionar entre ellos, estableciendo una temperatura del gas de unos 10.000 grados Kelvin (*d*).



UNA NEBULOSA GASEOSA TAMBIEN EMITE RADIACION cuando los electrones libres pasan cerca de los protones u otros núcleos cargados positivamente sin combinarse con ellos. En un encuentro rasante muy próximo de este tipo, el electrón pierde una considerable cantidad de energía, emitiendo en el proceso un fotón de radiación óptica de alta frecuencia. Sin embargo, la mayor parte de los electrones apenas son desviados, y emiten radioondas de frecuencia comparativamente baja. El resultado neto de un gran número de tales encuentros rasantes es la emisión de un espectro continuo de radiación que se transmite a través del espacio sin sufrir atenuación alguna por parte de los detritos interestelares. La radiación que se halla producida por este mecanismo se conoce como bremsstrahlung o "radiación de enfrenamiento".

es pequeño comparado con la duración de la vida de una nebulosa típica (alrededor de 10 millones de años), el número de recombinaciones por unidad de volumen es igual al correspondiente número de ionizaciones. Hay dos procesos físicos definidos que compiten aquí. Por una parte, la velocidad a la que un fotón ultravioleta interacciona con un elemento neutro próximo depende sólo de las propiedades atómicas del átomo del elemento y del número de fotones suficientemente energéticos, y es independiente de las propiedades termodinámicas del gas. Por otra parte, la velocidad de recombinación radiativa de un electrón con un ion es una función fuertemente dependiente de la temperatura y densidad del gas de la nebulosa.

En primera aproximación, la solución de las ecuaciones de equilibrio de ionización es notablemente sencilla. La solución, que especifica la estructura de la ionización y la geometría de la nebulosa, requiere que el flujo de radiación estelar se diluya a cierta distancia de la estrella central excitante, en parte porque algunos de los fotones ultravioleta se están usando sin cesar para ionizar nuevos elementos (en particular hidrógeno) a medida que la zona ionizada se expande, con rapidez al principio y más lentamente después, invadiendo la región neutra circundante. Después de unos 10.000 años o menos del nacimiento de la estrella central excitadora (o estrellas), el tamaño y la forma de la nebulosa alcanzan virtualmente un régimen estacionario, haciéndose casi por entero ionizada dentro de una región bien definida y quedando casi totalmente neutra más allá de ella. Tales regiones se llaman algunas veces esferas de Strömgren, en honor del astrónomo Bengt Strömgren, que fue el primero en resolver las ecuaciones del equilibrio de ionización.

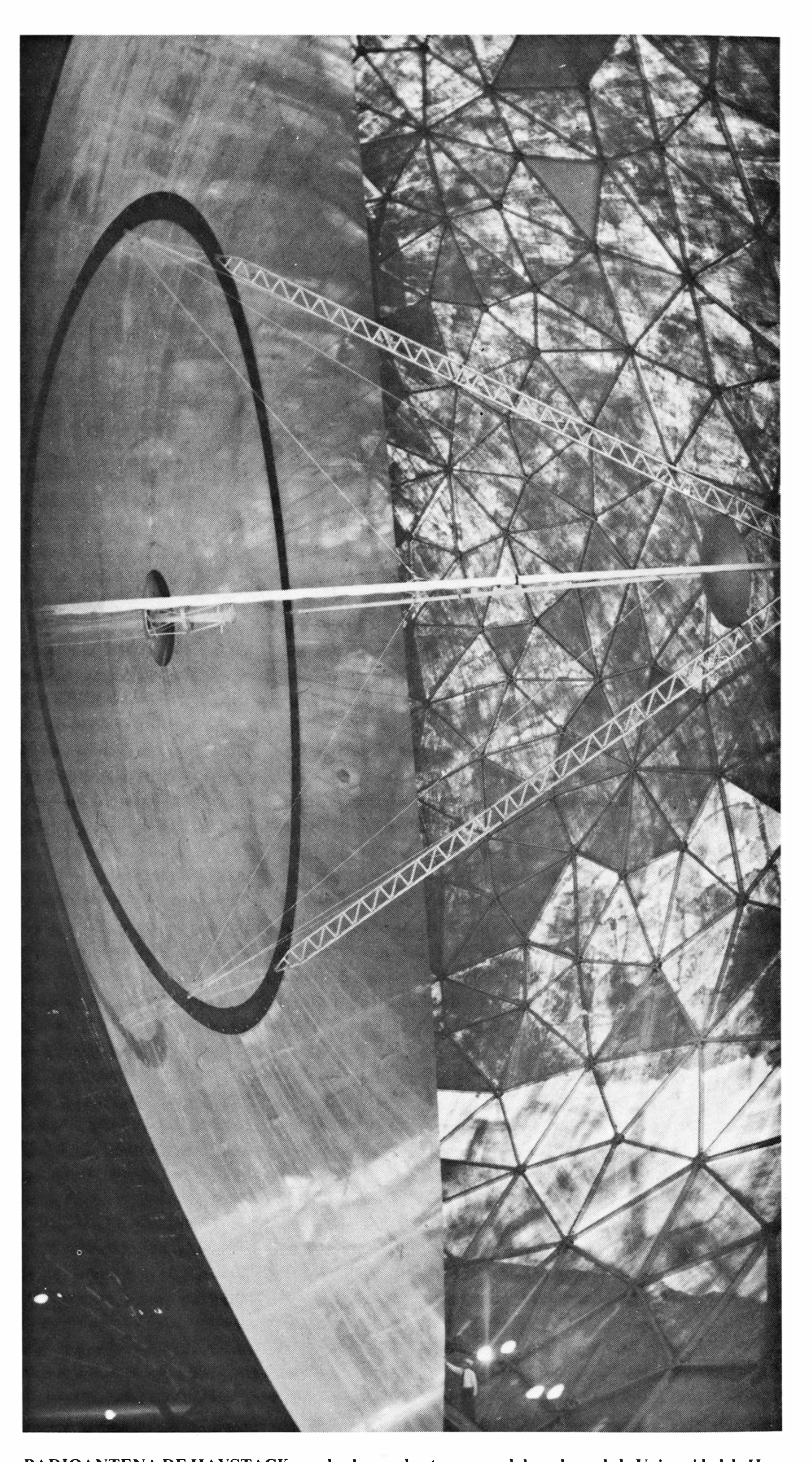
Las estrellas de distintos tipos espectrales tienen radios de Strömgren característicos. El cálculo del radio nebular depende poco de la temperatura que se suponga a la nebulosa, ya que la región ionizada está termostáticamente controlada por el ritmo de choques de los fotoelectrones con los iones escasos. Se puede esperar, sin embargo, que la densidad de la nebulosa varíe en más de un orden de magnitud respecto a un valor medio de 300 iones por centímetro cúbico. Por ejemplo, las nebulosas muy difusas (aquellas con una densidad de unos 10 iones por centímetro cúbico) se pueden extender hasta unos 20 años luz desde una estrella muy caliente. Pueden existir nebulosas todavía más grandes

cuando son varias estrellas muy calientes (o un verdadero cúmulo de estrellas) las que suministran colectivamente un gran flujo de fotones ultravioleta.

A pesar de la completa ionización de las regiones interestelares cercanas a las estrellas calientes, cada estrella o cúmulo de estrellas puede producir sólo un número finito de fotones ultravioleta, y, por tanto, sólo puede ionizar un volumen finito de materia. Naturalmente, si en la proximidad de una estrella caliente hay ausencia de materia susceptible de ser ionizada, entonces el tamaño de tal nebulosa de densidad acotada viene limitado por la cantidad de materia neutra presente, y su forma, generalmente irregular, depende de la distribución geométrica de la materia interestelar en la vecindad de la estrella. En fecha reciente, sin embargo, los astrónomos han podido darse cuenta de que la mayor parte de las nebulosas están inmersas en nubes mucho más grandes y mucho más masivas de átomos y moléculas neutras que rodean a las nebulosas "radiativamente encerradas". Tales nebulosas presentan, de ordinario, alguna clase de simetría alrededor de sus fuentes excitadoras. Las ecuaciones del equilibrio de ionización especifican además que el espesor de la región de transición entre el gas caliente y completamente ionizado de dentro de la nebulosa y el gas frío y predominantemente neutro de fuera de ella es extremadamente acusada en comparación con las dimensiones generales de la nebulosa. Sin embargo, dentro de esta región de transición, cuyo espesor se mide en horas-luz más bien que en años-luz, la temperatura y el grado de ionización del gas caen bruscamente, en varios órdenes de magnitud en un caso típico.

Una nebulosa gaseosa es, pues, una región caliente y casi por completo ionizada en su interior, separada por una capa de transición extremadamente delgada de una nube interestelar predominantemente fría y neutra. La información acerca de las nebulosas gaseosas se ha ido acumulando gracias, exclusivamente, a la detección de la radiación electromagnética emitida y absorbida por los electrones, iones y átomos neutros. Los astrofísicos clasifican esos datos electromagnéticos según cuatro variables fundamentales, a saber: dirección, frecuencia, intensidad y polarización, a fin de caracterizar el estado físico y químico del gas en función de la temperatura, densidad, composición y velocidad.

Como ejemplo típico, consideremos una región del espacio al nordeste de la constelación de Sagitario. Una fotografía con gran angular de esta parte del cielo



RADIOANTENA DE HAYSTACK, empleada por el autor y sus colaboradores de la Universidad de Harvard para detectar radioondas procedentes de las nebulosas gaseosas. La instalación, de cuyo funcionamiento la Northeast Radio Observatory Corporation ha encargado al Instituto de Tecnología de Massachusetts, está situada sobre una colina en Westford, Mass. Está encerrada dentro de una gran cúpula geodésica para protegerla de los inviernos de Nueva Inglaterra. En esta fotografía de "ojo de pez", hecha desde dentro de la cúpula, la superficie de la antena primaria de 37 metros aparece a la izquierda. La superficie primaria (que es extremadamente lisa, desviándose de un paraboloide matemáticamente perfecto en menos de medio milímetro en cualquier punto de su superficie) sirve para focalizar las radioondas cósmicas sobre una superficie secundaria más pequeña *(derecha)*, que a su vez refleja las señales hacia un receptor sensible situado en el centro del disco primario. El margen de funcionamiento de la antena está comprendido normalmente entre uno y 25 gigahertz; se ha usado para detectar radioseñales de frecuencia más alta.

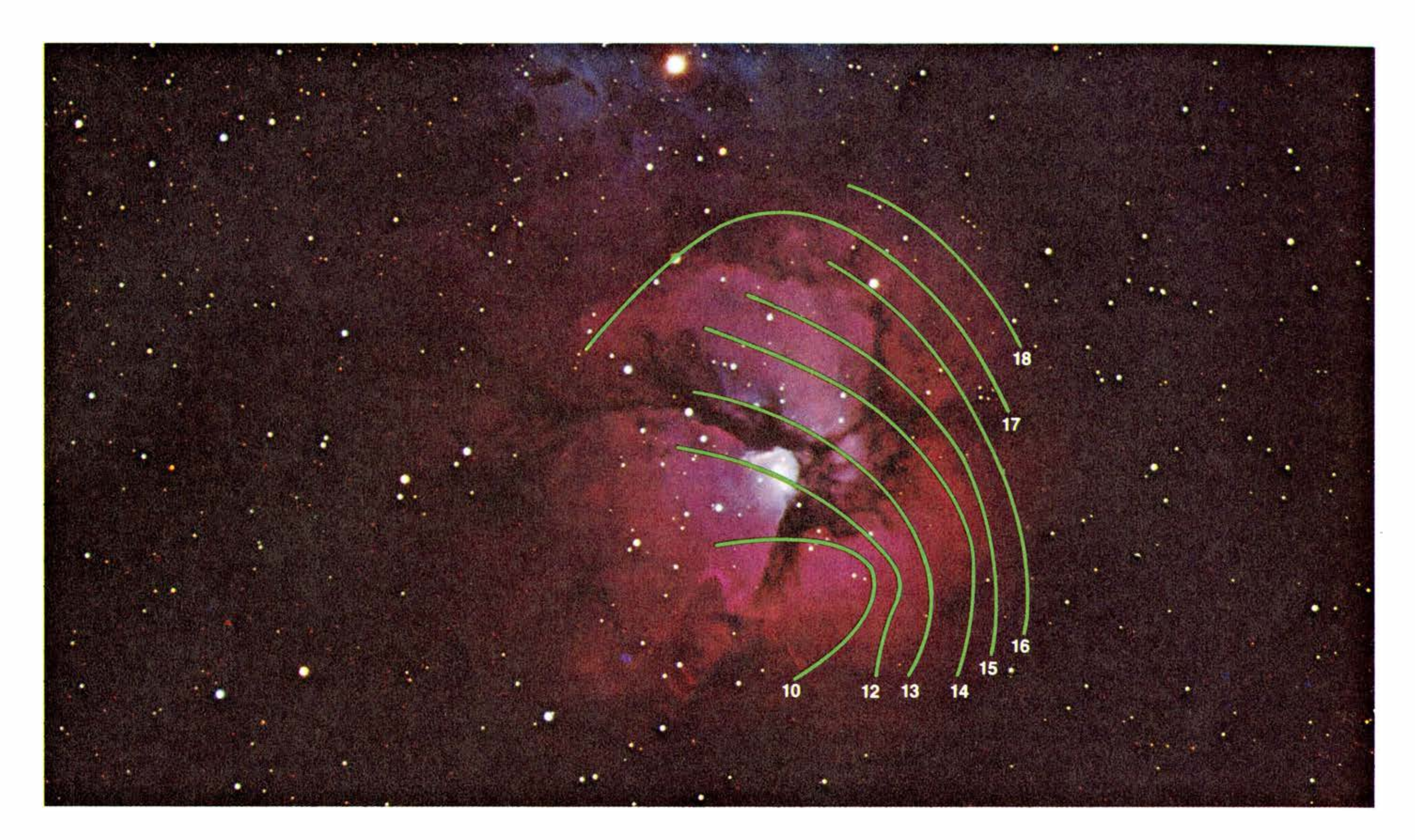
abarca el plano central de nuestra galaxia [*véase la ilustración de la página 83*]. Un examen cuidadoso de la copia positiva en blanco y negro revela una apariencia moteada, indicando una distribución irregular de estrellas. La aparente irregularidad es el producto de un obscurecimiento visual que enmascara lo que debería ser una distribución uniforme de estrellas. El obscurecimiento es particularmente apreciable a lo largo del plano galáctico mismo, donde la acumulación del polvo y otras formas de detritos interestelares es muy pronunciada. En realidad, el grado de obscurecimiento es tan alto que impide ver más allá de 10.000 años luz en esta dirección; por consiguiente, no hay posibilidad de usar telescopios ópticos para estudiar el núcleo de la galaxia, que se encuentra en la misma dirección a unos 30.000 años luz del sistema solar.

Inmersas en este vasto mar de polvo y gas interestelares se encuentran varias nebulosas gaseosas, cuatro de las cuales aparecen en color en las ilustraciones que se acompañan. Estas son algunas de las nebulosas gaseosas más llamativas que pueden observarse en parte alguna del cielo. Sus números de Messier (asignados a ellas por Charles Messier, el astrónomo francés del siglo XVIII que fue el primero en catalogarlas), junto con sus nombres populares, son M18, Nebulosa de la Laguna; M16, Nebulosa del Aguila; M17, Nebulosa Omega, y M20, Nebulosa Trífida. Una observación cuidadosa de esas nebulosas individuales muestra obscurecimientos adicionales, con frecuencia directamente a través de la imagen fotográfica. Esta extensa difusión de la luz por el polvo interestelar ayuda a dificultar el análisis de los datos ópticos. De hecho, algunas de las nebulosas más energéticas son completamente invisibles, en particular aquellas situadas hacia el centro de la galaxia.

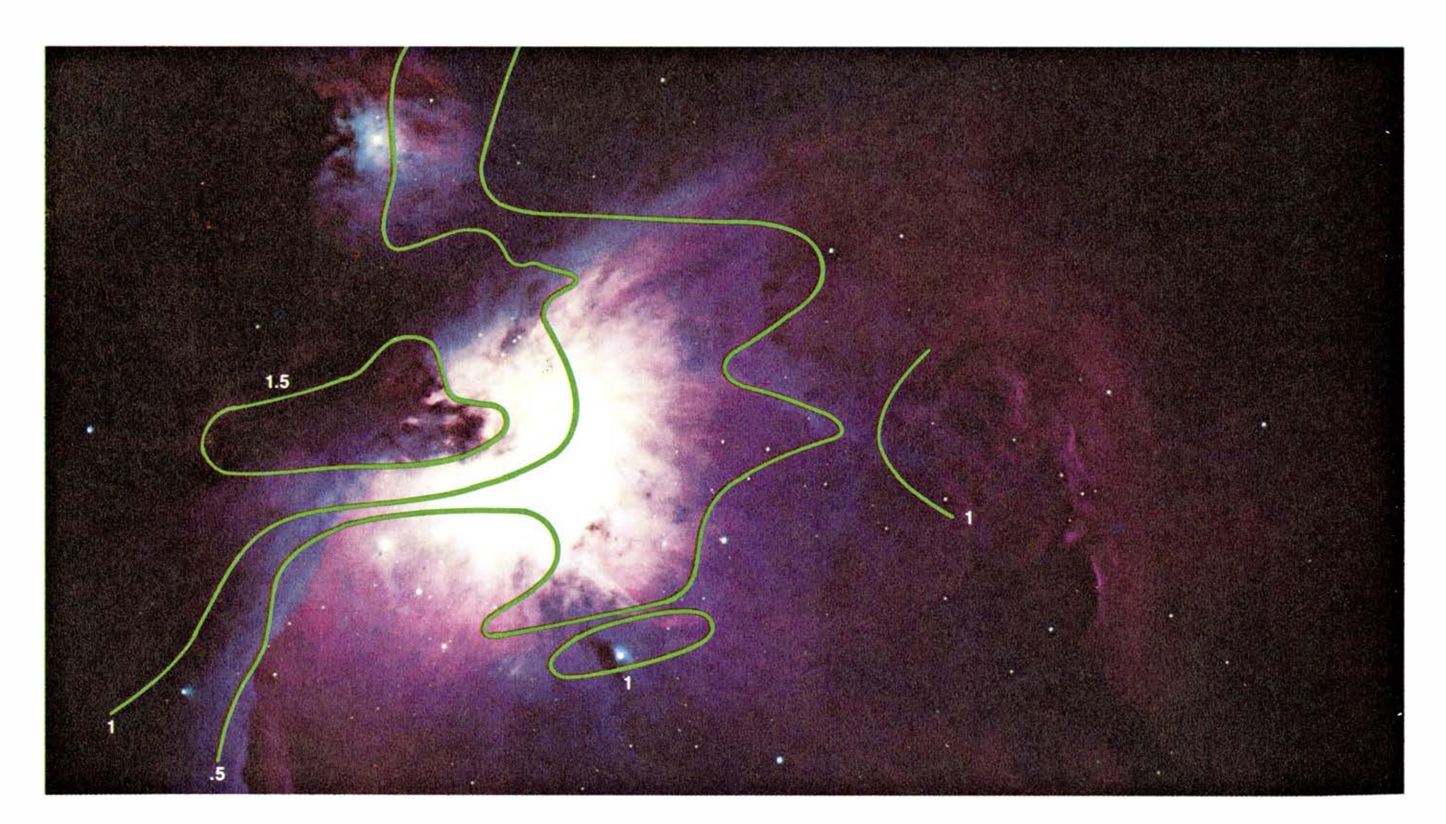
¿Cómo se puede obtener entonces una verdadera delineación fidedigna del gas en una nebulosa? Para ello hay que hacer uso del hecho de que las nebulosas emiten ondas de radio que no se atenúan en absoluto al atravesar el polvo interestelar. Consideremos de nuevo el papel que juegan los electrones libres. Durante los varios cientos de años que esperan para recombinarse con un núcleo, los electrones no ligados pierden parte de su energía cinética por medio de choques inelásticos con otras partículas cargadas. En un encuentro típico electrón-ion, el electrón, más ligero, es frenado al penetrar en el campo electrostático del ion, que suele ser un protón. Una interacción rasante de este tipo no conduce a la recombinación porque bien la velocidad de aproximación o la separación en el impacto es demasiado grande para que se produzca la captura. Para conservar energía, la pequeña cantidad de energía cinética perdida por el electrón se emite en forma de radiación de frecuencia comparativamente pequeña, esto es, ondas de radio, fenómeno que se conoce con el nombre de bremsstrahlung, palabra alemana que significa radiación de enfrenamiento o de frenado. Pese a la rareza de los encuentros individuales, el conjunto de los aproximadamente 10^{58} electrones en una nebulosa típica está dotado de una distribución continua de velocidades de aproximación y separación en el impacto, lo que garantiza la emisión continua de ondas de radio capaces de escapar de la nebulosa.

La débil emisión de bremsstrahlung suministra un poderoso método para delinear la verdadera distribución del gas nebular con un radiotelescopio grande. Consideremos, por ejemplo, un mapa de las radioondas emitidas por M20, la Nebulosa Trífida [*véase la ilustración en la parte superior de la página 81*]. Esta radiación se exploró en la frecuencia de ocho gigahertz (ocho mil millones de ciclos por segundo), correspondiente al dominio de las microondas, por la antena de Haystack de la Northeast Radio Observatory Corporation, un consorcio de las universidades del nordeste de los Estados Unidos. La resolución angular efectiva del telescopio en este caso era de casi cuatro minutos de arco. El radiomapa demuestra que con la excepción de las bandas trisecantes de material opaco, la imagen óptica de M20 es una representación razonablemente buena de la extensión del gas ionizado. Esto,

DESCRIPCION MECANICO-CUANTICA de un átomo de hidrógeno; se compone de un conjunto de niveles discretos de energía que está permitido ocupar al único electrón del átomo. Las transiciones del electrón entre los niveles de energía más bajos, representadas a la izquierda, dan lugar a la emisión de la luz visible. Las transiciones entre niveles más altos y menos espaciados, muy ampliadas en el centro, van acompañadas de la emisión de radiación en frecuencias de radioondas. Tres de tales transiciones se indican en color, así como el aspecto real de sus correspondientes líneas de radio-recombinación *(derecha)*, registradas en la dirección de la nebulosa M20. Las gráficas de las transiciones 111 a 110 y 73 a 72 se hicieron con ayuda de la antena de 43 metros del Observatorio Nacional de Radioastronomía. La curva de la transición 93 a 92 se hizo con la antena de Haystack. Las líneas características de radio-recombinación de este tipo quedan determinadas por tres mediciones: su intensidad, su anchura y la frecuencia de su centroide.



EN EL CASO DE LA NEBULOSA M20, este mapa de distribución de velocidades en el gas ionizado sugiere una lenta rotación de la nebulosa. El mapa se obtuvo analizando los centroides de cierto número de líneas de radio-recombinación. Las isolíneas muestran los distintos corrimientos Doppler de diversas partes de la nebulosa, medidas en kilómetros por segundo respecto a un sistema de referencia basado en un grupo local de estrellas cercanas al Sol. El mapa cinemático resultante, superpuesto a la misma fotografía de M20 de Kitt Peak reproducida en la página 81, parte superior, sugiere que la nebulosa gira alrededor de un eje orientado 45 grados hacia el Este contando desde el Norte con un período de varios millones de años. Caben otras interpretaciones.



ATENUACION DE LA LUZ de una nebulosa por el polvo interestelar y otros detritos galácticos. En el caso de M42, mostrada aquí, el polvo parece estar concentrado en la región obscura al este del núcleo de la nebulosa. La luz visible de esta zona comprendida dentro de la isolínea rotulada 1,5 ha sido atenuada casi en un 80 por cien; la luz dentro de la curva 1 es aproximadamente 60 por cien menos intensa de lo que sería en ausencia de polvo, y la luz dentro de la isolínea 0,5 sufre una atenuación de un 40 por cien. No se sabe si el polvo está físicamente mezclado con el gas nebular o si se encuentra en la forma de nubes de fondo no relacionadas con la nebulosa. La sorprendente asimetría de la nebulosa visible y la distribución no uniforme de las estrellas vecinas fueron las primeras indicaciones de que el volumen ionizado de la nebulosa gaseosa es sólo una pequeña parte de una agregación de materia interestelar neutra.

sin embargo resulta ser una circunstancia rara.

La Gran Nebulosa de Orión (M42) suministra un ejemplo mejor de cómo porciones importantes de gas nebular pueden estar visualmente ocultadas [*véase la ilustración inferior de la página 81*]. El mapa de radiointensidad observado en la dirección de M42 es más detallado porque a la frecuencia de operación de 24 gigahertz de esta observación, la respuesta angular era aproximadamente un minuto de arco. La resolución es todavía pobre comparada con la que se obtiene mediante telescopios ópticos, pero es la mejor que actualmente se puede obtener con un radiotelescopio. Dos o más radiotelescopios pueden funcionar sincrónicamente para formar un interferómetro capaz de resolución superior, pero esos dispositivos son sensibles sólo a una pequeña fracción del gas nebular.

Los radiomapas de resolución comparativamente alta, tales como el de M42 se pueden usar conjuntamente con estudios ópticos para formular una descripción más cuantitativa acerca de la ocultación interpuesta. La radioondas, que no son en absoluto atenuadas por el polvo interpuesto, se pueden multiplicar en frecuencia en varios órdenes de magnitud a fin de obtener la intensidad verdadera de la radiación óptica. La diferencia entre los valores calculados (no atenuados) y observados (atenuados) de la intensidad óptica proporciona una medida del enturbiamiento.

Este método se empleó recientemente comparando el radiomapa de M42 que yo hice con el instrumento de Haystack con un mapa óptico de M42 hecho por M. A. Dopita, de la Universidad de Manchester, en el observatorio de alta montaña del Pic du Midi, en los Pirineos. La distribución resultante del enturbiamiento de fondo se puede presentar en silueta frente al gas brillante de M42 [*véase la ilustración inferior de la página anterior*]. Como podría esperarse, la obscura lengua de polvo hacia el este del núcleo de la nebulosa corresponde a la región de máximo enturbiamiento.

Estudios análogos han mostrado que porciones considerables de algunas nebulosas, en particular M17, son completamente invisibles. De hecho, cientos de nebulosas gaseosas invisibles para los astrónomos ópticos se han descubierto en los últimos años mediante técnicas radioastronómicas.

Además de proporcionar una medida del grado de enturbiamiento, los radiomapas de radiación nebular continua pueden suministrar una gran riqueza de información acerca del estado termodinámico del gas nebular. Por ejemplo, la intensidad total de las radioondas emitidas depende evidentemente del número de electrones que chocan con iones dentro de un volumen dado del espacio. Se puede invertir entonces la descripción matemática del proceso de bremsstrahlung para determinar la densidad electrónica o iónica, conocida la radiointensidad observada. Si se conoce la distancia de la nebulosa, su tamaño físico se puede calcular de su tamaño aparente; la masa de la nebulosa se puede determinar a partir de su intensidad y volumen conocidos. En la tabla de la página 94 se relacionan las estimas de distancia, radio de Strömgren, densidad y masa de las cinco nebulosas prominentes en este artículo. Las densidades y masas dadas dependen algo de un modelo teórico, pero no obstante representan satisfactoriamente las propiedades generales de cientos de nebulosas distribuidas en toda la galaxia. Comparando el tamaño observado de una nebulosa con su radio de Strömgren teórico, se puede incluso deducir el tipo espectral de la estrella o estrellas centrales excitadoras. De esta manera se puede estudiar la distribución energética de la nebulosa sin siquiera tener que ver la estrella central excitadora.

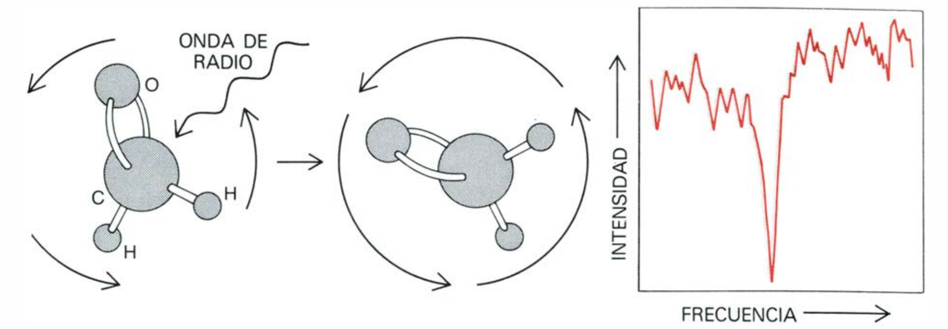
Con el tiempo, los electrones libres se recombinan con protones y a veces con núcleos más complicados existentes en el gas nebular. La física cuántica predice que los electrones capturados pueden tener sólo energías discretas, correspondientes a los niveles energéticos atómicos descritos más arriba. Como los electrones ligados buscan el nivel de energía más bajo, las sucesivas transiciones hacia abajo liberan paquetes de ondas (fotones) de energía determinada y cuantificada, que dan lugar a líneas espectrales claramente definidas y superpuestas, en frecuencias precisas, al espectro de fondo de radiación continua. El estudio de esas características espectrales permite al astrónomo examinar cualquier especie atómica individual en la nebulosa, con exclusión de todas las otras. Tales análisis espectroscópicos de nebulosas continúan suministrando datos experimentales muy precisos, con los cuales se pueden construir modelos más parecidos a la realidad de las nebulosas y su entorno interestelar.

La cascada descendente de un electrón en la escalera cuantificada de niveles de energía ligados conduce primero a emisión de ondas de radio procedentes

PERFIL DE INTENSIDAD DE LAS RADIOONDAS procedentes de M42, deducido de los mismos datos que el mapa topográfico de la nebulosa presentado en la ilustración inferior de la página 81. En este caso, sin embargo, la distribución de la radiointensidad se observa desde la perspectiva de un ángulo de elevación de 30 grados. Esta clase de representaciones, registradas en una gran pantalla de televisión en Haystack, ponen de relieve la mayor radiointensidad en el centro de la nebulosa y muestran cómo las nebulosas gaseosas y otras fuertes radiofuentes destacan del "ruido" de la radiación galáctica de fondo.

de niveles altos en el átomo, después a radiación infrarroja y, finalmente, a radiación visible. La frecuencia de la emisión varía porque la diferencia de energía entre niveles atómicos adyacentes se hace más pequeña al crecer la distancia al núcleo central. Por ejemplo, las líneas visibles permitidas del hidrógeno y del helio, tal como se ven con un espectroscopio óptico, son en realidad recombinaciones de líneas que llevan consigo transiciones electrónicas entre los niveles más bajos de energía del átomo. Sus intensidades están con frecuencia enmascaradas por la difusión, de manera parecida a como lo está la radiación continua de la luz de las estrellas.

Solamente ahora se están construyendo espectrómetros de infrarrojo para observar las líneas de recombinación de más baja frecuencia procedentes de transiciones electrónicas entre niveles atómicos superiores al quinto. Los espectrómetros de radiofrecuencia incorporados a los grandes radiotelescopios pueden, sin embargo, detectar fácilmente líneas de recombinación de muy baja frecuencia procedentes de transiciones electrónicas entre niveles de energía superiores al trigésimo. De hecho, las líneas de radio-recombinación procedentes de nebulosas gaseosas se están ahora observando habitualmente en frecuencias que varían desde 300 megahertz a 100 gigahertz. En las frecuencias más bajas, los electrones que saltan en cascada entre niveles atómicos excepcionalmente altos, a veces aproximándose al nivel 300, permiten a los radioastrónomos estudiar los mayores átomos jamás observados espectroscópicamente. Esos átomos gigantes tienen un diámetro de unas cinco milésimas de milímetros y no se pueden estudiar en los laboratorios terrestres porque en las condiciones, densas en comparación, de incluso un laboratorio de "vacío", los choques despueblan los electrones de los niveles atómicos altamente excitados incluso antes de que tengan oportunidad de emitir radiación.



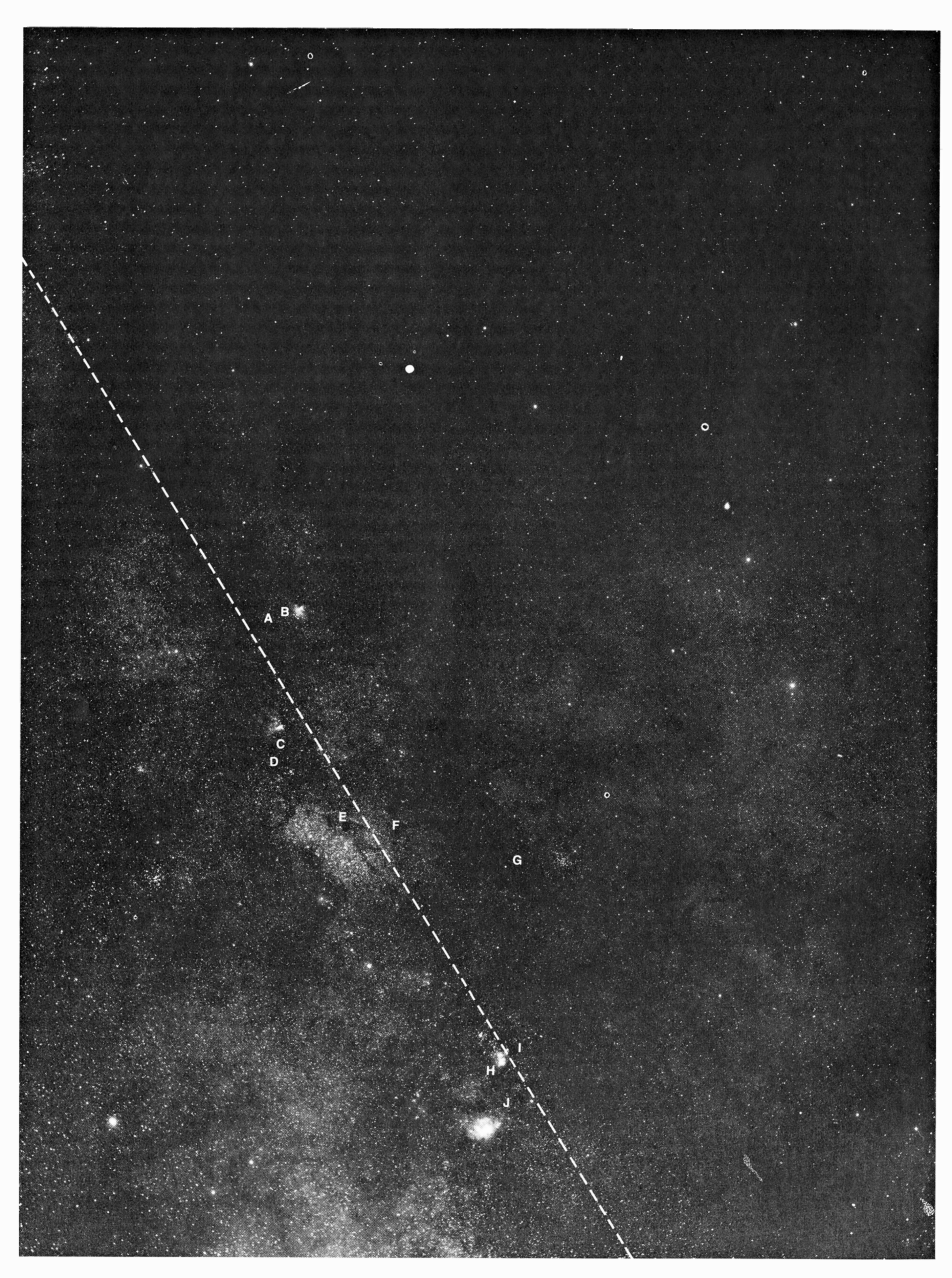
LAS MOLECULAS DE FORMALDEHIDO en el espacio interestelar giran a velocidades específicamente cuantificadas según su contenido de energía. Los cambios en la configuración rotacional de una tal molécula van acompañados de emisión o absorción de radiación de radiofrecuencia, que se puede detectar en la Tierra en forma de una figura espectral característica tal como la que aparece a la derecha. En el par de diagramas a la izquierda, un fotón de frecuencia comparativamente baja (de radiofrecuencia) emitido por una nebulosa se ve cómo está siendo absorbido por una molécula de formaldehído (H_2CO) en una nube de fondo, lo que produce que la molécula de formaldehído gire más velozmente. La característica espectral particular, ilustrada a la derecha, fue registrada en M20 mediante la radioantena de 43 metros del NRAO.

Consideremos la apariencia física de una línea espectral de radiofrecuencia a medida que los electrones nebulares descienden en cascada del nivel 111 al nivel 110 en el átomo de hidrógeno [*véase la ilustración de la página 82*]. Esta observación se realizó hacia M20 con el espectrómetro multicanal unido al radiotelescopio de 42,5 metros del Observatorio Nacional de Radioastronomía en Green Bank, Virginia Occidental. Este tipo de líneas en forma de campana se especifican matemáticamente por su intensidad, su anchura y la frecuencia de su centroide. Naturalmente, tales características espectrales provienen no de un único átomo, sino de un gran número de electrones que sufren una transición determinada dentro del volumen nebular explorado por el telescopio.

Las líneas espectrales pueden suministrar generalmente un gran volumen de información acerca del gas emisor o absorbente, sea éste una muestra de laboratorio o una muestra cósmica. Mis colaboradores de Harvard y yo en el Observatorio Haystack usamos las líneas de radio-recombinación para extraer información física, química y cinética de las medidas de características espectrales de distintos elementos en una gran variedad de lugares diferentes. Así, hemos construido un mapa de isolíneas de la velocidad del gas en M20 [*véase la ilustración superior de la página 87*]. Este mapa se dedujo de observaciones de una línea de radio-recombinación realizadas en muchos puntos distribuidos en la nebulosa. La frecuencia central de la línea espectral se convirtió en velocidad radial usando el efecto Doppler, según el cual, el cambio de frecuencia de la radiación se puede utilizar para interpretar el movimiento de acercamiento o alejamiento. Debe recalcarse que el movimiento Doppler tiene lugar sólo a lo largo de la línea de observación; los movimientos transversales de objetos distantes no pueden estudiarse por este método.

Un examen detenido del mapa de las velocidades del gas sugiere que esta nebulosa está girando alrededor de un eje situado aproximadamente a 45 grados al este del norte. Si tal es el caso, dada la distancia a M20, el gradiente de velocidad implica una rotación como un cuerpo sólido, con un período de algunos millones de años. Esto representa bastante lentitud; de hecho, el período de rotación es mayor que la edad que se sospechaba para una nebulosa joven de polvo tal como M20. Hay también otras interpretaciones posibles; en particular, las observaciones pueden ser el resultado de la dilatación diferencial, o cizalladura, del gas nebular caliente contra el entorno circundante frío. Se necesitarán observaciones futuras, hechas con instrumentos capaces de una resolución angular más fina, para desentrañar las numerosas ambigüedades tales como ésta, referentes a la dinámica global del gas nebular.

El análisis cuidadoso de un línea espectral puede proporcionar mucha información adicional acerca de las condiciones físicas en la nebulosa. Consideremos los diversos mecanismos que pueden ensanchar una línea espectral. El calor es una fuente obvia de ensanchamiento de líneas, ya que cada uno de los electrones que experimenta una transición determinada hacia abajo está dotado de un movimiento aleatorio de agitación; algunos movimientos térmicos son de alejamiento, otros de aproximación, pero todos sirven para ensanchar la línea a causa del efecto Doppler. Los movimientos turbulentos, no térmicos, de la materia nebular en bloque en una escala que es pequeña comparada con la resolución angular del telescopio son otra fuente del ensanchamiento de líneas; las células microturbulentas de gas tienen distintos movimientos, y cuando se promedian para la totalidad de la nebulosa, pueden ensanchar o deformar apreciablemente el perfil espectral. Los movimientos del gas en gran escala, en particular la rotación, podrían contribuir también al ensanchamiento de alguna característica observada, si bien ninguna de ellas ha sido observada sin ambigüedad. Además, los choques inelásticos de átomos que se recombinan con el plasma, particularmente con los electrones libres, pueden ensanchar de



RADIOOBSERVACIONES RECIENTES de formaldehído en la región del espacio alrededor de la dirección general del centro galáctico delinean una vasta nube de moléculas neutras que al parecer engloba muchas nebulosas, incluyendo M8, M16, M17 y M20. Esta fotografía de gran ángulo es la misma que la representada en la página 83. Las letras dan la posición de las medidas de líneas de formaldehído *(véanse gráficas de la página siguiente).*

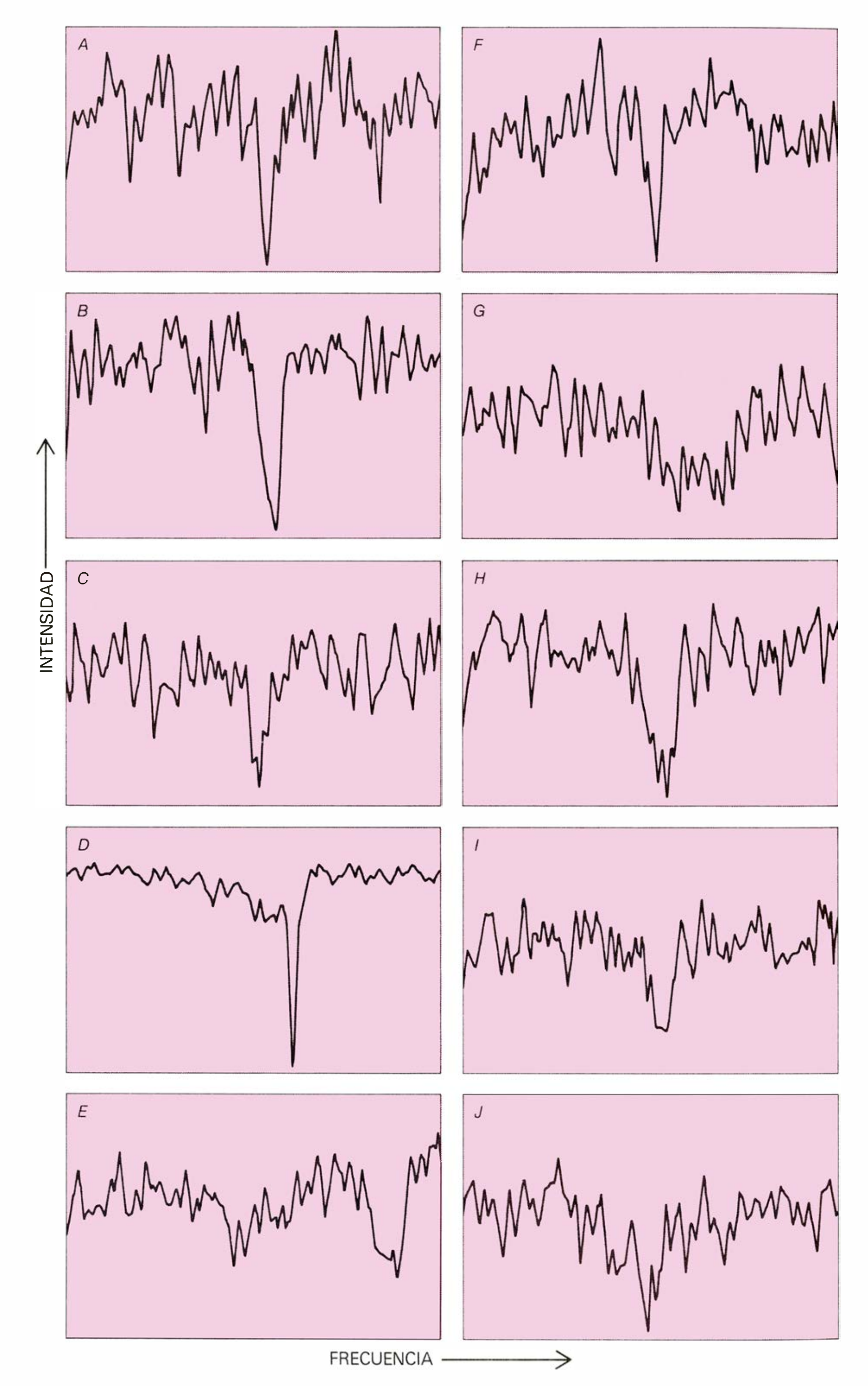
un modo apreciable una línea espectral, sobre todo si los electrones ligados que sufren el choque se encuentran suficientemente alejados de su núcleo en el momento de la colisión.

Por fortuna para nosotros, la formulación matemática de muchos de esos mecanismos de ensanchamiento de líneas se puede realizar en forma precisa, de modo que es posible especificar con exactitud los factores físicos dominantes en el ensanchamiento de líneas espectrales cuando éste está asociado a determinados fenómenos de recombinación. Por ejemplo, los movimientos Doppler térmicos de los átomos individuales son evidentemente función de la temperatura del gas: cuanto más caliente sea el gas, tanto mayor será el movimiento atómico microscópico y tanto más ancho será el perfil de la línea. Análogamente, un gas denso garantiza choques más frecuentes y, por tanto, una línea espectral más ancha.

El objetivo del análisis de las características espectrales observadas es invertir el conocimiento teórico que poseemos de los mecanismos individuales de ensanchamiento de líneas, desentrañar la magnitud de las diversas contribuciones y extraer información razonable acerca de la nebulosa. Esta tarea se ha realizado con éxito para varias nebulosas observando muchas líneas de recombinación del mismo elemento a frecuencias diferentes, y algunas veces de diferentes elementos a casi la misma frecuencia. Los valores de la temperatura y densidad del gas hallados de esta manera concuerdan notablemente bien con los obtenidos del análisis del continuo de radiación en radiofrecuencia. Los valores de la turbulencia gaseosa son, en general, menores que la velocidad esperada para el sonido, de 11 kilómetros por segundo en un tenue gas a 10.000 grados Kelvin. Sin embargo, algunas nebulosas, en particular aquellas cuyas exigencias de excitación son altas, como la M17, muestran algunos indicios de turbulencia supersónica, que puede considerarse bien como el producto de frentes de choque nebulares o como el movimiento acumulado de células individuales subsónicas mantenidas por vientos de alta velocidad, procedentes de varias estrellas excitadoras.

A partir de observaciones de las líneas de radio-recombinación pueden determinarse también las abundancias químicas. En particular, hemos obtenido valores para la relación de iones de helio a iones de hidrógeno para varias nebulosas. El valor medio, de aproximadamente 9 por ciento, es igual a la relación de abundancias helio a hidrógeno para el cosmos si, como debe esperarse de consideraciones de la energética de las nebulosas, no hay cantidades apreciables de helio neutro o doblemente ionizado.

En realidad, las determinaciones de abundancia de helio son de considerable importancia, no sólo para mejorar nuestra comprensión de las nebulosas gaseosas, sino para la totalidad de la ciencia astrofísica. Como el ciclo protón-protón que origina helio por fusión en los interiores estelares da cuenta generalmente de una relación helio/hidrógeno de menos del uno por cien, dada la duración de la vida de la galaxia, los cosmólogos generalmente tratan de explicar



REGISTROS de la línea de absorción del formaldehído en las diversas posiciones indicadas por las letras de la fotografía de la página opuesta, que muestran que la absorción es máxima en las regiones más polvorientas, donde las condiciones son más adecuadas para la formación y protección de moléculas frágiles. Los registros fueron obtenidos por el autor y sus colaboradores con la antena del Observatorio Nacional.

la producción del restante 8 o 9 por cien de helio atribuyéndolo a las primeras fases de la "gran explosión" que se cree originó el universo. Los objetos para los cuales la relación de helio a hidrógeno es substancialmente inferior al 10 por cien de la abundancia cósmica presentan, por tanto, dificultades potenciales para muchas teorías cosmológicas, ya que no hay manera de destruir el helio una vez se ha formado. Así, las medidas de la abundancia de helio hechas con ondas de radio no dispersadas, que tienen su origen en un gran número de objetos galácticos, pueden en último término tener implicaciones profundas para la comprensión de la naturaleza del propio universo.

En resumen, las intensidades, anchuras y centroides de las líneas espectrales pueden cartografiarse en toda la extensión de las nebulosas. De esta manera, las variaciones angulares de temperatura, densidad, turbulencia, composición y velocidad han podido obtenerse para diversas nebulosas en un esfuerzo para aumentar nuestro conocimiento sobre las propiedades del plasma en la vecindad de estrellas jóvenes.

La explicación anterior representa sólo una breve exploración de un conjunto bastante detallado de parámetros físicos o que se pueden obtener de estudios espectroscópicos de nebulosas gaseosas. Hay, naturalmente, algunas hipótesis críticas inherentes a la interpretación de las líneas espectrales nebulares. Sin embargo, la frontera del tema está en gran parte relacionada con mediciones finas de características espectrales de varios elementos en un intento de apreciar la validez de esas hipótesis básicas y de ahí determinar con precisión la naturaleza del plasma interestelar. De hecho, algunos investigadores que trabajan en este campo han llegado a alcanzar la sutileza de discutir sobre si la temperatura del gas nebular es de 8000 o 9500 grados K, y sobre si la densidad del gas es de 500 o de 800 electrones por centímetro cúbico. Ello representa una precisión notable en astronomía galáctica. Está en marcado contraste con muchos campos de la astrofísica en los que las especulaciones teóricas, por no hablar de las observaciones, no se pueden hacer con precisión superior a un orden de magnitud. Es realmente una indicación de la confianza que los radioastrónomos tienen de alcanzar pronto una comprensión completa de la materia en una de sus formas no terrestres más comunes.

Para adquirir una comprensión completa de las nebulosas gaseosas, sin embargo, se debe examinar también su entorno interestelar circundante, porque allí es donde la materia galáctica se condensa en primer lugar para formar las estrellas calientes. Hasta hace poco tiempo se creía que las nebulosas gaseo-

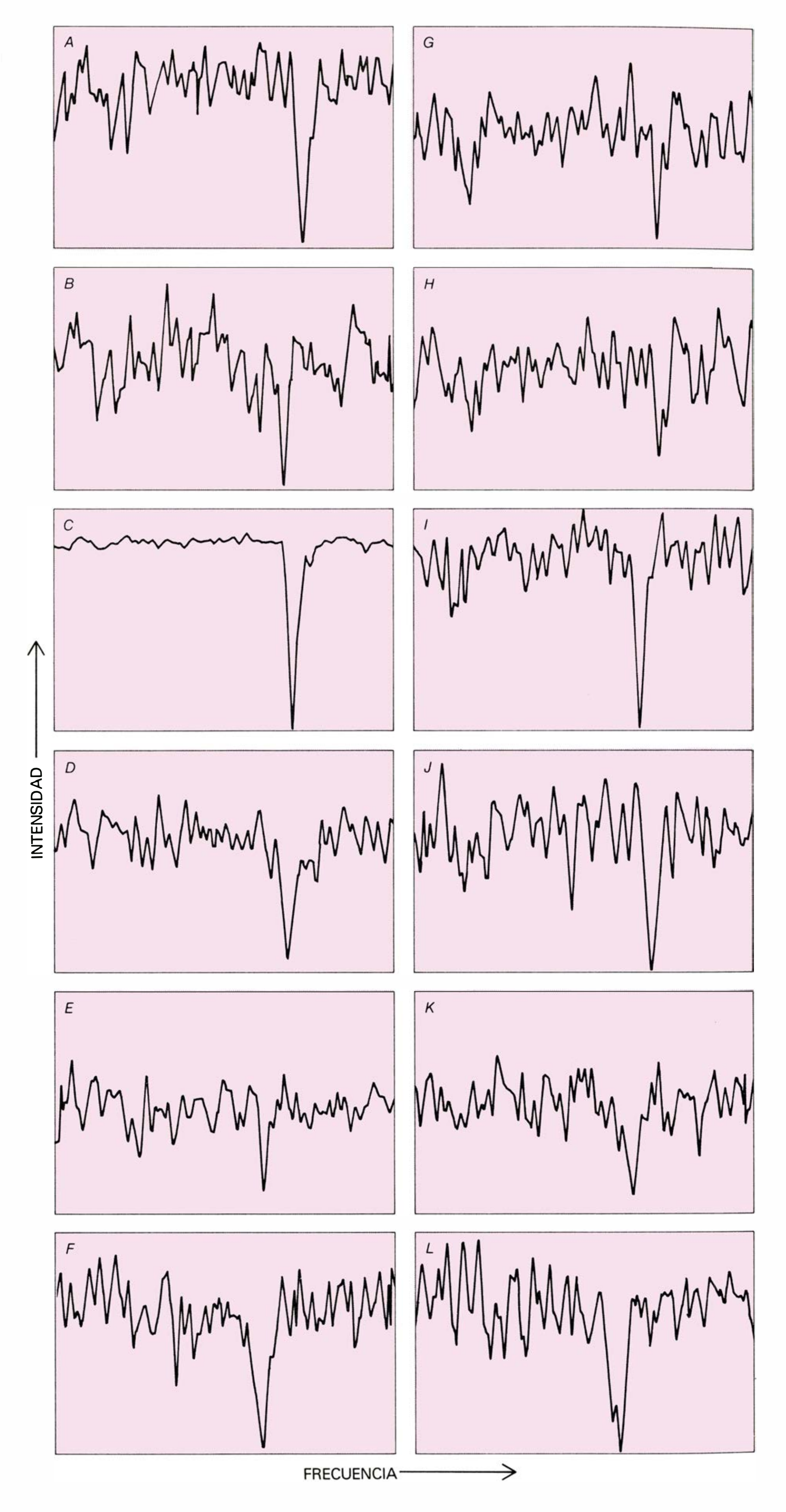


LA CONSTELACION DE ORION también parece estar englobada en una vasta nube molecular, según medidas de la línea de absorción característica del formaldehído en varios puntos de esta región del cielo. La fotografía, que mide 13 por 19 grados, se realizó con una de las cámaras de vigilancia del cielo de 7,6 centímetros en una de las estaciones septentrionales de observación de Harvard en Massachusetts. La nebulosa de Orión (M42) propiamente dicha es la amplia mancha de aspecto nebuloso que rodea a la estrella central de las tres estrellas verticamente alineadas en el centro. En las representaciones clásicas, esas tres estrellas corresponden a la espada de Orión, que cuelga de su cinto, definido por la línea de tres estrellas más brillantes inmediatamente encima. Las letras se refieren a los registros que se recogen en la página siguiente.

sas se contaban entre los objetos más masivos de la galaxia. Desde hace muy pocos años, sin embargo, los radioastrónomos han venido descubriendo indicaciones de la existencia de enormes y densas nubes moleculares que dejan enanas a las nebulosas gaseosas tanto en tamaño como en masa. Muchas de esas nubes moleculares se encuentran en los brazos espirales de la galaxia, regiones que son ricas en gas y polvo, pero están desprovistas de nebulosas. Sin embargo, se sabe que varias de las más prominentes nubes están íntimamente asociadas con nebulosas gaseosas. En efecto, se halla ahora muy extendida la hipótesis de que las nubes moleculares desempeñan un importante papel en la secuencia de la evolución preestelar de la materia interestelar. Por consiguiente, el conocimiento completo de las nebulosas gaseosas incluye necesariamente la descripción de su heterogéneo entorno galáctico.

Hay varias maneras de estudiar la materia neutra en la vecindad de una nebulosa. Se ha indicado más arriba una técnica de observación usada para estimar la opacidad a lo largo de la línea de visión hasta la nebulosa. Pueden recogerse indicaciones adicionales sobre los detritos interestelares explorando visualmente el campo no uniforme de estrellas adyacentes a una nebulosa. Numerosas zonas obscuras ocultan con frecuencia, parcial o totalmente, una distribución de estrellas de fondo por lo demás uniforme. Por ejemplo, casi todas las nebulosas visuales ilustradas aquí quedan encerradas en una zona con bajo recuento de estrellas; en particular, se encuentran tales zonas al sur y oeste de M17 y también de M20. Tales heterogeneidades en la distribución de estrellas, incluyendo la aparición de varias manchas pequeñas que producen ocultación completa, se hacen todavía más patentes en fotografías de exposición larga. Otra disciplina en la que hay considerable potencial para dilucidar el carácter de la materia neutra en la vecindad de las nebulosas es la radioastronomía de infrarrojo. En el momento presente, los estudios de infrarrojo implican claramente la presencia de polvo, aunque no se sabe si éste está mezclado con el gas nebular o es exterior a él. En realidad, todavía es virtualmente imposible determinar la relación física precisa entre las regiones polvorientas de materia que enturbia y las regiones ionizadas de gas nebular.

En todo caso, estos métodos presentan obvias limitaciones. La comparación entre las radiaciones óptica y de radio se puede hacer únicamente para las escasas regiones en que la radiación se puede



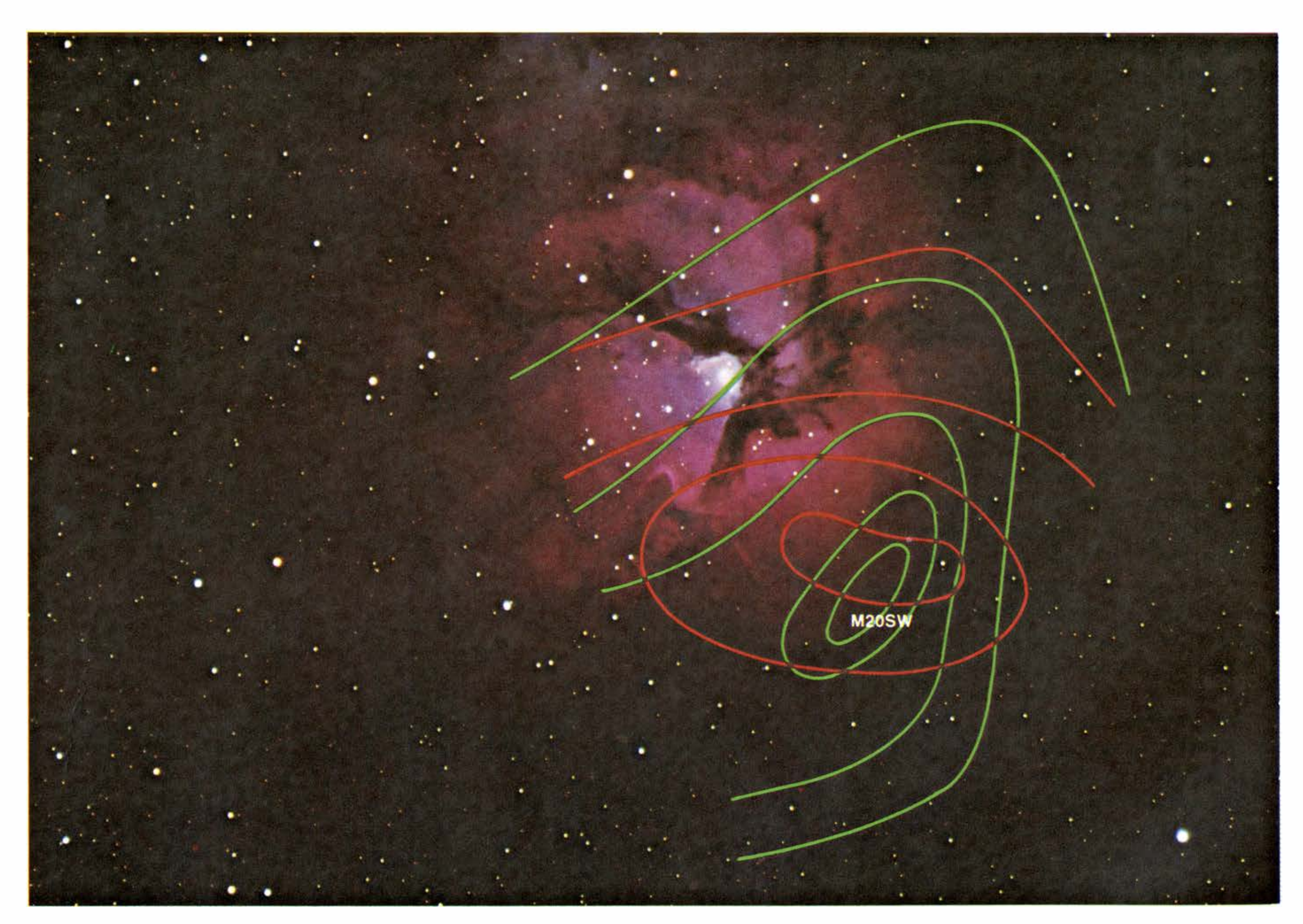
FIGURAS ESPECTRALES características de la absorción de radioondas por formaldehído interestelar, registradas en 12 puntos designados por las letras en la fotografía de la constelación de Orión en la página anterior. Todas las medidas coinciden con nubes obscuras, que sirven para proteger las moléculas contra su destrucción por la radiación ultravioleta procedente de las estrellas más próximas. (Albert Miller.)

NOMBRE DE LA NEBULOSA	NUMERO DE MESSIER	DISTANCIA DE LA TIERRA (AÑOS-LUZ)	RADIO DE STRÖMGREN (AÑOS-LUZ)	DENSIDAD (ELECTRONES POR CENTIMETRO CUBICO)	MASA (MULTIPLOS DE LA MASA SOLAR)	TEMPERATURA DE LOS ELECTRONES (GRADOS KELVIN)	VELOCIDAD DEL GAS TURBULENTO (KILOMETROS POR SEGUNDO)	PORCENTAJE He^+/H^+
LAGUNA	M8	3500	4	400	60	7500	9	9
AGUILA	M16	8000	12	100	600	8000	12	6
OMEGA	M17	7000	5	800	300	8700	17	10
TRIFIDA	M20	6500	8	100	150	8200	7	10
ORION	M42	1500	1	2000	10	9000	10	9

RELACION DE LAS PROPIEDADES de las cinco nebulosas representativas a las que se hace referencia en este artículo. El radio de Strömgren, mide la región del espacio dentro de la cual el gas nebular está ionizado por la radiación ultravioleta procedente de la estrella central. La última columna da la relación de iones helio (He^+) a iones hidrógeno (H^+) en las nebulosas; el valor medio (alrededor del 9 por cien) es igual a la relación cósmica He^+/H^+.

detectar visualmente; los recuentos de estrellas sólo ofrecen mediana exactitud, aun para aquellos investigadores decididos a emprender tan laborioso trabajo, y la astronomía de infrarrojo carece hasta ahora de técnicas espectroscópicas y de conocimiento adecuado de las propiedades emisivas de las partículas de polvo. Parece que la técnica de observación de momento mejor adaptada para enfocar el problema del entorno nebular es la espectroscopía en radiofrecuencia de las moléculas interestelares.

La astronomía molecular es una nueva ciencia interdisciplinaria que está en rápido desarrollo. En los últimos ocho años se han detectado, por medio de la espectroscopía de microondas y ondas milimétricas, más de 40 nuevas especies moleculares y 100 nuevas líneas espectrales. Algunas de esas moléculas, tales como el radical hidroxilo (OH) y el formaldehído (H_2CO), absorben radiación de intensas radiofuentes de



EL FORMALDEHIDO ESTA CONCENTRADO en la proximidad de una región obscura al suroeste del centro de M20, la Nebulosa Trífida. Las isolíneas verdes, que indican aumento de la radiointensidad en incrementos uniformes desde el exterior hacia el interior, muestran que la concentración de formaldehído alcanza un máximo en la zona rotulada M20SW. Las isolíneas color naranja, que también aumentan en incrementos uniformes desde el exterior hacia el interior, representan la distribución de velocidades del formaldehído en términos del ensanchamiento Doppler de la línea característica de absorción del formaldehído. Como sólo el movimiento a lo largo de la visual produce tal efecto Doppler, el mayor ensanchamiento de la línea en M20SW implica que la nube puede encontrarse en estado de colapso gravitatorio en ese punto. Los datos pueden así indicar el inminente nacimiento de una estrella.

fondo. Otras, como el monóxido de carbono (CO), monosulfuro de carbono (CS), cianuro de hidrógeno (HCN) y alcohol metílico (CH_3OH), emiten radiación propia. Todavía hay otros, tales como el agua (H_2O) y el monóxido de silicio (SiO), que emiten radiación intensa como consecuencia a un proceso análogo al del máser.

A pesar de la novedad de la astronomía molecular, la presencia de moléculas bastante complicadas en el espacio interestelar no es muy sorprendente. Las propiedades físicas y químicas básicas de más de un millón de especies moleculares diferentes se han ido acumulando en estudios de laboratorio durante años, pero sólo hace muy poco que los astrofísicos se dieron cuenta de que ciertas moléculas son virtualmente ubicuas en toda la galaxia. La mayor parte de los procesos químicos que sirven para formar moléculas se han estudiado sólo bajo condiciones terrestres, y de aquí que no se conozcan bien en el espacio interestelar, donde las densidades suelen ser varios órdenes de magnitud menores que los mejores vacios alcanzables en los laboratorios terrestres. Sin embargo, la idea de que incluso las moléculas sencillas serían fácilmente destruidas por la radiación interestelar ha sido abandonada al hacerse patente que grandes cantidades de polvo no sólo proporcionan lugar adecuado para la formación de moléculas, sino que también sirven para protegerlas de la asoladora radiación del espacio interestelar. En realidad, en los últimos años se ha consagrado como un hecho de observación el de que la abundancia de todas las moléculas está fuertemente correlacionada con regiones interestelares de fuerte enturbiamiento.

¿Cómo se pueden detectar las moléculas en el medio interestelar? Una técnica utiliza las modificaciones en su configuración electrónica, en analogía con las transiciones electrónicas descritas más arriba para los átomos. Esas alteraciones, sin embargo, conducen generalmente a radiación ultravioleta de alta frecuencia que es difundida por los detritos interestelares y no se puede observar desde la Tierra porque la atmósfera terrestre la bloquea por completo. En cambio, el movimiento en bloque de las moléculas mismas generalmente da lugar a líneas espectrales observables. Las moléculas giran sobre sí mismas y vibran como un muelle enrollado, simplemente en virtud de su calor. Esto lo hacen, sin embargo, a ritmos precisos de acuerdo con las reglas de la mecánica cuántica. Los modos de vibración suelen dar lugar a líneas espectrales características en el dominio relativamente inaccesible del infrarrojo. Por otra parte, una amplia variedad de moléculas, desde sencillas moléculas diatómicas inorgánicas hasta las complicadas moléculas orgánicas poliatómicas, giran a frecuencias específicas capaces de ser detectadas en el espectro de radiofrecuencia.

Varios procesos que compiten entre sí sirven para excitar las moléculas y así poblar estados rotacionales específicos de una manera muy parecida a como los estados de alta energía de los átomos se pueblan mediante los efectos radiativo y de choque. La sucesión más corriente de hechos tiene lugar como sigue. Fuentes estelares locales de radiación ultravioleta calientan el polvo, que mediante choques cede energía al hidrógeno molecular (H_2), la especie molecular neutra más abundante, que a su vez excita a moléculas más complicadas por medio de choques. El calor recibido lo es, por tanto, primariamente en forma de radiación estelar, mientras que el mecanismo de enfriamiento dominante se cree que resulta de la emisión de radiación que acompaña a las transiciones entre niveles de alta energía de rotación de las moléculas de monóxido de carbono. Las densas nubes moleculares, sin embargo, se hacen frecuentemente opacas a esta radiación del monóxido de carbono, creando un desequilibrio en el balance de calentamiento y enfriamiento. A falta de un mecanismo eficaz de enfriamiento, ciertas partes de la nube (las más densas) pueden entonces empezar a calentarse. Están en camino de convertirse en estrellas.

La ubicua molécula de formaldehído sirve para ilustrar algunas de las consideraciones precedentes. Absorbiendo o emitiendo radiación, esta especie molecular poliatómica puede experimentar cambios en su velocidad de rotación a la frecuencia característica (entre otras) de 4830 megahertz. Esa frecuencia corresponde a una variación en la configuración rotacional de la molécula de casi cinco mil millones de veces por segundo. Los modernos espectrómetros de radiofrecuencia pueden detectar con facilidad tal radiación como una absorción o emisión característica a una longitud de onda de unos seis centímetros, no difundida por los detritos interestelares ni atenuada tampoco por la atmósfera terrestre.

Las líneas moleculares de este tipo se pueden analizar de una manera muy parecida a como se hace con las líneas atómicas. Los perfiles de líneas moleculares generalmente suministran valiosa información sobre la temperatura, la densi-

dad, la composición y los movimientos de la nube molecular. Además, el conocimiento preciso de la física de la molécula bajo observación y de la velocidad de actuación del proceso de choque descrito más arriba proporciona una estima de la densidad de la especie dominante en el choque: el hidrógeno diatómico neutro. De esta manera, puede deducirse la densidad y masa totales de una nube molecular, supuesto que la extensión de la nube esté completamente determinada. En general, las nubes moleculares son frías, densas y masivas, tan masivas, en realidad, que la mayor parte de ellas dejan enanas a sus nebulosas gaseosas asociadas.

En una dramática ilustración de cuán grandes pueden ser las nubes moleculares, detecté recientemente moléculas de formaldehído en numerosos lugares entre M16 y M8, una separación de más de 11 grados en el cielo. (La luna llena sólo tiene medio grado de diámetro La velocidad casi constante de esas características sugiere que hay una nube gigante de materia molecular extendida desde M16 a M17, desde M17 a M20 y posiblemente desde M20 a M8. Ello implica que cada una de esas nebulosas se formó a partir de lo que era, y en su mayor parte todavía podría ser una única nube interestelar increíblemente vasta.

A la distancia media de las nebulosas (unos 7000 años luz) la nube tendría una dimensión de unos 1300 años luz, al menos en el plano del cielo. Con certeza no es una nube homogénea, porque el formaldehído se puede detectar sólo en aquellos lugares que son particularmente obscuros y opacos; esto es en sí una notable comprobación de la fuerte correlación entre la abundancia de moléculas y la cantidad de polvo. La masa de esta vasta nube, que no está todavía completamente delineada, podría ser más de 10 millones de veces la masa del Sol. Una nube molecular de semejante extensión encierra al parecer la nebulosa M42 y gran parte de la constelación Orión. Nuestras recientes observaciones en esta parte del cielo muestran de nuevo que dondequiera que haya una región obscura hay formaldehído, y por implicación, que hay igualmente otras moléculas.

Es evidente que nuestra galaxia está ricamente poblada de una gran diversidad de moléculas, principalmente asociadas a nebulosas gaseosas, aunque algunos objetos infrarrojos y nubes obscuras desprovistas de nebulosas abrigan números substanciales de moléculas. Las pequeñas diferencias cinemáticas (de movimiento) observadas en las velocidades atómicas (nebulares) y moleculares (interestelares) no son sorprendentes, porque si realmente las nebulosas se forman a partir del colapso de una condensación molecular, lo harán predominantemente en la periferia de la nube, ya que el 90 por cien de la masa de la esfera uniformemente densa está concentrado en su mitad exterior. Así, parece probable que la mayor parte de las nebulosas visibles, en particular los objetos considerados aquí, representan los fragmentos ionizados de la parte anterior de nubes moleculares más grandes.

¿Hasta qué punto encajan estas observaciones en el esquema a grandes rasgos de la formación de estrellas sugerido al comienzo de este artículo? Consideremos una vez más el complejo M20 como ejemplo representativo de una región potencialmente rica en estrellas recién nacidas. Un mapa de isolíneas de concentración de formaldehído en la proximidad de M20 parece englobar esta nebulosa gaseosa por el sur y el oeste [*véase la ilustración de la página 94*], imitando la distribución de opacidad por polvo. De hecho, los máximos de abundancia molecular coinciden con una mancha obscura totalmente opaca que hemos rotulado M20SW. Hay menos moléculas al este y nordeste de la mancha, donde la densidad de estrellas de fondo aparece menos atenuada. De manera interesante, la nebulosa de reflexión azulada al nordeste del complejo M20 parece contener suficiente polvo para difundir la luz de las estrellas (de igual manera que la luz del Sol se difunde en la atmósfera terrestre para dar al cielo su color azul), pero no el suficiente para proteger las moléculas de su destrucción por la radiación.

La misma ilustración incluye un mapa de la anchura de la línea del formaldehído, que también tiene un máximo en M20SW. La simetría común de la distribución cinemática y de la concentración molecular sugiere que la condensación molecular en M20SW puede encontrarse en estado de colapso gravitatorio; en la periferia de la nube se observan características moleculares más estrechas porque el movimiento de materia que cae hacia dentro es predominantemente transversal a la línea de visión y no influye en el ensanchamiento del perfil de la línea por efecto Doppler. Si esta interpretación es correcta, el complejo M20 incluye uno de los primeros ejemplos conocidos que aportan pruebas de observación al esquema de evolución interestelar que he descrito aquí.

Una revisión de nuestros hallazgos en el complejo M20 parece suministrar indicaciones provisionales de la existencia de tres fases principales en la formación de estrellas. En el noroeste del complejo, la velocidad nebular de 18 kilómetros por segundo, deducida del análisis de líneas de radio-recombinación, concuerda muy bien con la velocidad de la nube molecular adyacente, estableciendo con ello la fase inicial o tranquila, en la cual la densidad es probablemente no muy alta (quizá 100 partículas por centímetro cúbico). La velocidad de 18 kilómetros por segundo puede considerarse como una medida del movimiento de la nube interestelar gigante antes de la formación de la estrella excitadora de M20. En ciertos lugares, por ejemplo en M20SW, la nube molecular quizás ha alcanzado una densidad suficientemente alta (mayor de 1000 partículas por centímetro cúbico) y una masa lo bastante grande (mayor de 2000 masas solares) para iniciar el colapso gravitatorio, o fase intermedia de la formación de estrellas.

La propia nebulosa M20 representa la fase final de la formación de estrellas. La curva más exterior de nuestro radiomapa de la nebulosa a ocho gigahertz muestra dos irregularidades que bien podrían ser reales. La primera es un saliente hacia el nordeste, donde las fotografías muestran alguna ionización y donde la nube molecular está ausente o al menos es tenue. La segunda es un entrante en el límite nebular más próximo a M20SW. Es fácil visualizar una onda de ionización creada por la acción de la radiación estelar ultravioleta propagándose en la densa región molecular en la proximidad de M20SW. El frente de choque que precede a la ionización de la materia ciertamente comprimiría el gas y probablemente contribuiría al incremento de densidad en la nube molecular cerca de M20SW. Es incluso concebible que tal onda de choque pudiera actuar como mecanismo desencadenante de una rápida formación de estrellas en esa posición.

Las futuras observaciones de las radioondas emitidas por diversas especies atómicas y moleculares, en particular observaciones con mejor resolución espacial, pueden hacer posible la estimación cuantitativa del grado de interacción inducida por la onda de choque entre el caliente y tenue gas nebular y la densa y fría materia molecular. Puede que entonces sea posible identificar una fase adicional intermedia entre la fase de colapso y la fase nebular, a saber: una fragmentación a pequeña escala de la nube en lugares individuales de formación de estrellas.

Mecanismo molecular de la transducción de energía

El conocimiento de la subestructura de las ATPasas traslocadoras de protones permite avanzar en la resolución, a nivel molecular, de un problema básico en biología: la transformación de la energía química en formas asimilables

Emilio Muñoz y J. Manuel Andreu

La vida de los organismos depende en gran parte de un eficaz aporte de energía por y para sus células. La falta de energía conduce a las células a la muerte, ya que, incapaces de crecer, dividirse y de moverse para la captura de nutrientes, terminan por agotar sus propias reservas. Tres son los procesos fundamentales utilizados por los seres vivos para producir energía: la glucolisis, la fosforilación oxidativa y la fotofosforilación. Los tres procesos requieren sistemas multienzimáticos eficazmente coordinados, pero la fosforilación oxidativa y la fotofosforilación exigen, además, que esos sistemas estén localizados en sistemas membranosos. Las células eucarióticas llevan a cabo estos procesos en orgánulos especializados: las mitocondrias, donde se efectúa la fosforilación oxidativa, y los cloroplastos, donde tiene lugar la fotofosforilación. Las células procarióticas disponen de estos sistemas enzimáticos integrados en su membrana plasmática. En estos tres tipos de membrana (mitocondrial, cloroplástica y plasmática) se lleva a cabo el proceso de transducción de energía, acoplando la energía derivada del transporte de electrones a la síntesis del trifosfato de adenosina (ATP). El proceso global de la ruta de síntesis de ATP comprende tres fases o pasos: la captura de energía en los complejos de oxidación, la transmisión de energía desde estos complejos a los complejos de fosforilación propiamente dichos y la utilización de la energía en estos complejos para sintetizar el trifosfato de adenosina.

El mecanismo de estos procesos es una cuestión que ha preocupado mucho durante los últimos 15 años. Slater propuso en 1953 la llamada hipótesis química, que postulaba una analogía entre los procesos de fosforilación oxidativa y fotofosforilación con la fosforilación a nivel de substrato, preconizando la existencia de intermediarios químicos para cada uno de los pasos mencionados. Numerosos grupos de investigadores han dedicado denodados esfuerzos en las últimas décadas a la detección de estos intermediarios con resultados negativos. Recientemente, esta hipótesis, ya muy desacreditada, ha vuelto a ganar un cierto predicamento merced a los trabajos de Griffths y colaboradores. Estos investigadores han demostrado que, en la síntesis de ATP, pueden intervenir intermediarios químicos, como oleil-S-ácido lipoico y oleil-fosfato, ya que el ácido dihidrolipoico actúa como un componente del complejo encargado de la síntesis de ATP.

A consecuencia de la serie de resultados negativos en la búsqueda de intermediarios, se pensó en la necesidad de elaborar hipótesis alternativas. En 1964, Boyer proponía su hipótesis de acoplamiento energético conformacional que, en un sentido amplio, postulaba que todos y cada uno de los tres procesos citados anteriormente debían depender de los cambios conformacionales de las proteínas integrantes de los complejos enzimáticos, cambios que serían, a su vez, dependientes del aporte de energía y que se transmitirían por interacciones proteína-proteína dentro del espacio membranal. Sin embargo, a lo largo de estos 15 años transcurridos no se han presentado pruebas experimentales de peso específico suficiente, que apoyen la idea de que la captura y transmisión de energía tiene lugar por una red proteica interconectada. Merced a las ideas y al trabajo pionero de Peter Mitchell se desarrollaba otra alternativa, la hipótesis quimiosmótica, que es la que goza de mayor predicamento actualmente entre los investigadores del campo de la bioenergética. Para explicar la transmisión de energía, en su más amplio sentido, la hipótesis quimiosmótica considera que la distribución asimétrica de los componentes de las cadenas de óxido-reducción da origen a potenciales eléctricos transmembranales y/o gradientes de *p*H, a través de los cuales se transmite y se utiliza la energía para la síntesis de ATP. A lo largo de estos últimos años se han acumulado una cantidad considerable de pruebas experimentales que apoyan la relación entre cambios en el potencial transmembranal y/o en el gradiente de *p*H con la síntesis o hidrólisis de ATP (véase "Cómo fabrican ATP las células", por Peter C. Hinkle y Richard E. McCarty, INVESTIGACION y CIENCIA, mayo, 1978).

La hipotesis quimiosmótica parece explicar el mecanismo de la transmisión de energía desde unos complejos a otros. Constituye la estructura conceptual para entender el estado energizado de las membranas, el proceso por el que las membranas biológicas favorecen la separación asimétrica de cargas, que determina a su vez el desarrollo de un gradiente de protones. Sin embargo, no explica el mecanismo molecular de la síntesis de ATP a partir del difosfato de adenosina (ADP) y fosfato inorgánico (P_i) gracias a la utilización de la energía transmitida. En un intento de explicar todos los pasos de la transducción de energía, Mitchell ha propuesto que la energía transmitida a los complejos de fosforilación es utilizada en éstos para sintetizar ATP por medio de protonaciones de un átomo de oxígeno del fosfato inorgánico. Sin embargo, no hay pruebas que apoyen este punto, que sigue siendo el más conflictivo y en el que la solución parece por el momento difícil.

Es conveniente resaltar que el interés por los conceptos quimiosmóticos ha

desviado la atención del problema más difícil de explicar químicamente en la transducción de energía, y que es, precisamente, la síntesis de ATP. El conocimiento de los mecanismos responsables de este proceso pasa por la caracterización de los componentes moleculares implicados en dicho proceso.

Las ATPasas traslocadoras de protones son elementos esenciales de la maquinaria transductora de energía, en la que desempeñan el papel esencial de intervenir en la formación de gradientes de H^+ y en la síntesis de ATP. Para cumplir con esta complejidad funcional, se organizan como sistemas muy complejos, constituidos por un mínimo de ocho polipéptidos organizados estructuralmente en, al menos, dos componentes supramoleculares: uno, de carácter hidrofóbico, se puede definir operacionalmente como constituyente intrínseco de membrana y se designa como F_0 (factor de acoplamiento 0) y otro de mayor talla, con un carácter más hidrofílico, que se designa como factor F_1 (factor de acoplamiento 1). El factor F_0 facilita el tránsito de protones a través del medio hidrofóbico de la membrana impermeable a los iones, mientras que el factor F_1, en virtud de su carácter de proteína periférica de membrana, se ha podido solubilizar por tratamientos suaves, sin necesidad de recurrir a detergentes. Estos procedimientos no implican, por tanto, desnaturalización de la proteína. Con ello se ha facilitado su purificación y su caracterización molecular. Se han purificado factores F_1 de diferentes orgánulos (mitocondrias de corazón de buey y de hígado de rata, cloroplastos de varias especies vegetales, cromatóforos de microorganismos fotosintéticos) y de membrana plasmática de diferentes procariotas.

A las ATPasas procedentes de mitocondrias se les designa genéricamente como factores F_1, mientras que para los cloroplastos y bacterias se utilizan las abreviaturas CF_1 y BF_1, respectivamente. La información sobre estos factores es considerable y permite abordar un estudio comparativo de sus propiedades que dé luz sobre la relación estructura-función en estas proteínas, sobre todo teniendo en cuenta las peculiaridades fisiológicas de su material de procedencia. En este sentido, el estudio de las ATPasas translocadoras de protones de origen bacteriano está contribuyendo de una forma muy decisiva a esclarecer el problema, sobre todo por la interesante propiedad que se le ofrece al experimentador para poder cambiar la batería enzimática de las bacterias por manipulación genética y fisiológica.

Todos los factores F_1 poseen propiedades, moleculares y funcionales comunes. Entre las primeras cabe destacar la existencia de masas moleculares muy semejantes, del orden de 320.000-350.000 dalton, y a la que contribuye un complejo patrón de subunidades. En este último aspecto hay que señalar la constante presencia de dos subunidades mayoritarias, referidas como α y β, cuyos pesos moleculares son asimismo relativamente homogéneos dentro de los diferentes factores F_1, ya que oscilan entre 58.000 ± 5000 dalton para α y 55.000 ± ± 3000 para β. Es interesante resaltar que la diferencia en peso molecular entre α y β para un determinado F_1 no es superior a 5000 dalton. Ello permite concluir que las dos subunidades mayoritarias poseen, efectivamente, pesos moleculares muy cercanos y que diferencias superiores observadas pueden responder a propiedades intrínsecas del factor F_1 estudiado y/o a la técnica utilizada por un determinado grupo de experimentadores. El carácter general de estas propiedades parece firmemente establecido. Sin embargo, conviene señalar que los valores de las masas moleculares pueden estar sujetos a revisión en un futuro, ya que su determinación, tal como se ha efectuado hasta ahora, presenta limitaciones, derivadas de la misma complejidad de los factores y de las técnicas utilizadas.

Por lo que respecta a las subunidades minoritarias, la situación es menos definida. Todos los factores F_1 y CF_1 aislados de mitocondrias y cloroplastos constan de tres subunidades minoritarias: una subunidad, en proporción constante, γ (peso molecular 30.000-35.000 dalton) y otras subunidades, δ y ε cuyos pesos moleculares oscilan alrededor de 17.000 y 12.000 dalton. La subunidad ε de CF_1 actúa como inhibidor natural de la actividad ATPasa de la proteína, mientras que en el caso de F_1 la identidad entre el inhibidor natural y la subunidad ε es más conflictiva. Para unos autores se cumple esta identidad, mientras que para otros el inhibidor natural de F_1 es una sexta subunidad de peso molecular intermedio entre las dos subunidades δ y ε. En el caso de las ATPasas bacterianas (factores BF_1) la situación con respecto a las subunidades minoritarias es menos coherente. Se han aislado ATPasas constituidas únicamente por α y β, las dos subunidades mayoritarias (*Bacillus megaterium*, formas *L* de *Proteus*). En otros factores BF_1, en particular los de especies del género *Micrococcus*, se ha detectado un número intermedio de subunidades minoritarias, que oscila de 1 a 3 con la presencia constante de una subunidad γ de peso molecular alrededor de 32.000. La presencia de una subunidad δ de peso molecular alrededor

HIPOTESIS	EL MECANISMO DE ACOPLAMIENTO				LA SINTESIS DE ATP	
	PRIMERA FASE	APOYO EXPERIMENTAL	SEGUNDA FASE	APOYO EXPERIMENTAL	TERCERA FASE	APOYO EXPERIMENTAL
QUIMICA	$AH_2 + B + C \rightleftarrows A \sim C + BH_2$	NO	$A \sim C + D \rightleftarrows C \sim D + A$	NO	$C \sim D + Pi + ADP \rightleftarrows C + D + ATP$	ALGO ?
CONFORMACIONAL	$AH_2 + B \rightleftarrows A^o + BH_2$	POCO	$A^* \rightleftarrows C^*$	MUY POCO	$C^* + Pi + ADP \rightleftarrows ATP + H_2O + C$	SI
QUIMIOSMOTICA	$2H^+$ / $AH_2 + B$ / $A + BH_2$	MUCHO	$2H^+$ $2H^+$ $2H^+$	BASTANTE	$Pi + ADP + 2H^+ \rightleftarrows ATP + H_2O + 2H^+$	?

LAS TRES FASES DE LA SINTESIS del trifosfato de adenosina (ATP) a partir del difosfato de adenosina (ADP) y el fosfato inorgánico (P_i) en relación con las tres hipótesis de la transducción de energía. La hipótesis química (Slater, 1953) postula la existencia de compuestos intermediarios "ricos en energía". La hipótesis conformacional (Boyer, 1964) supone que la energía derivada de la respiración y la fotosíntesis se acumula en cambios conformacionales de los complejos enzimáticos. La hipótesis quimiosmótica (Mitchell, 1961, 1966) introduce el concepto vectorial en las reacciones de óxido-reducción. Los transportadores de electrones se disponen de modo que la circulación de los electrones origina la translocación de protones a través de la membrana; si ésta no es permeable, se genera una diferencia de la actividad electroquímica de los protones entre el exterior y el interior.

ABREVIATURA	ORIGEN	PESO MOLECULAR (METODO)	SUBUNIDADES (LIMITES PESO MOLECULAR) α	β	γ
F_1	Mitocondrias	360.000-347.000 (equilibrio sedimentación) 384.000-360.000 (cromatografía en gel)	(62.500-53.000)	(57.000-49.000)	(39.000-25.000)
CF_1	Cloroplastos	325.000 (equilibrio sedimentación) 310.500 (calculado n.º subunidades y pesos moleculares)	(62.000-59.000)	(57.000-56.000)	(38.000-37.000)
BF_1	*Escherichia coli* Gram⁻, mesofilo, facultativo	290.000 (equilibrio sedimentación) 400.000-365.000 (cromatografía en gel)	(66.000-54.000)	(60.000-52.000)	(35.000-30.000)
BF_1	*Streptococcus faecalis* G⁺, mesofilo, anaerobio	350.000 (gradiente densidad) 385.000 (equilibrio sedimentación)	(60.000-55.000)	55.000-50.000)	35.000
BF_1	*Micrococcus lysodeikticus* G⁺, mesofilo, aerobio.	345.000-385.000 (equilibrio sedimentación diferentes formas)	(63.000-53.000)	(58.000-47.000)	(35.000-29.000)
BF_1	PS_3 G⁺, termofilo	380.000 (cromatografía en gel)	56.000	53.000	32.000

RESUMEN de las propiedades moleculares de ATPasa-factores F_1 purificados a partir de diferentes membranas (mitocondrial, cloroplástica o plasmática de basterias). Los factores F_1 son el componente, soluble en agua, de los complejos ATPasa-ATP sintetasa que tienen su asiento en las membranas capaces de llevar a cabo la transducción de energía. Sobre la base de esta propiedad, los factores F_1 han podido purificarse utilizando los métodos convencionales de purificación de proteínas y caracterización molecular se ha emprendido con gran interés y entusiasmo durante el último decenio. Estos estudios entrañan serias dificultades, a causa de la gran complejidad molecular de los factores F_1. La determinación de las masas moleculares de las F_1-ATPasas es un problema todavía no resuelto, si bien los resultados convergentes sugieren que los valores obtenidos son significativos. Nótese que en el caso de los sistemas F_1 bacterianos se recogen propiedades morfológicas y fisiológica de los organismos estudiados. La electroforesis en gel de poliacrilamida en condiciones desnaturalizantes (dodecil sulfato sódico o urea) ha sido la técnica de elección para determinar los tipos de subunida-

de 20.000 se detecta con menos frecuencia, mientras que componentes de baja talla que se podrían asimilar a la subunidad ε existen en la gran mayoría de las preparaciones de BF_1 de *Micrococcus lysodeikticus*, aunque su evidente falta de homogeneidad hace inviable su asimilación a un determinado tipo de subunidad. El factor BF_1 de *Escherichia coli* es el que posee una subestructura más compleja entre las F_1-ATPasas bacterianas, ya que en todos los casos en que se ha purificado la proteína se ha visto que está compuesta, al menos, de 4 subunidades, α, β, γ y ε, mientras que la presencia de δ es opcional según el procedimiento de purificación y/o análisis.

De lo que antecede, se deduce que los factores F_1 poseen una serie de elementos estructurales básicos comunes, que se han conservado a lo largo de la evolución. Sin embargo, también se observan diferencias que en unos casos son mínimas, pero en otros bastante acusadas. No se sabe si estas diferencias representan propiedades intrínsecas de las proteínas como reflejo de ciertas peculiaridades fisiológicas, o bien si son resultados artificiales derivados o amplificados por la manipulación experimental. Creemos que ambos hechos coexisten, como discutiremos con más detalle al referirnos a la flexibilidad conformacional de estas proteínas y a la posibilidad de aislar distintas formas de las ATPasas.

Un paso siguiente en la caracterización de los factores F_1-ATPasas consiste en la posibilidad de asignar una determinada función de la ATPasa a una subunidad definida. Se ha realizado un evidente progreso en este campo, aunque los resultados obtenidos encierran un carácter fragmentario por lo que la generalización y el carácter definitivo de las conclusiones están aún por ver.

Hay pruebas en favor del papel que la subunidad más pequeña, o una de las más pequeñas (con peso molecular alrededor de 10.000 dalton), juega, como inhibidor natural de la actividad hidrolítica para el ATP de los factores F_1, CF_1 y BF_1 de *E. coli*. La situación es menos clara en el caso de las ATPasas o factores BF_1 de bacterias gram-positivas, donde la presencia de estas subunidades, o no se han detectado (caso de BF_1 de *Bacillus megaterium*), o si se ha detectado no se ha puesto en evidencia su analogía funcional, como ha sido el caso de la ATPasa de *St. faecalis* o de una bacteria termofílica designada abreviadamente como PS_3 (Y. Kagawa y colaboradores, de la Universidad de Jichi, Japón).

En los estudios realizados en nuestro laboratorio sobre el factor BF_1 de *M. lysodeikticus*, hemos podido poner de manifiesto la existencia de un inhibidor natural de la actividad ATPasa que en este caso se identifica con una subunidad de peso molecular alrededor de 25.000 dalton que sería más asimilable al de δ, aunque su identidad no ha sido establecida de manera inequívoca. También se ha observado en nuestro laboratorio que la mayor estabilidad de la ATPasa de *M. lysodeikticus* a su digestión por tripsina está relacionada con la presencia de polipéptidos asociados, pero distintos de las subunidades más o menos definidas (α, β, γ, δ o componentes de movilidad relativa 1,0).

Las funciones de las subunidades de la ATPasa se han estudiado por experimentos con anticuerpos contra cada una de las subunidades, por digestión parcial de la ATPasa con enzimas proteolíticos, principalmente tripsina, o por experimentos de reensamblaje, a partir de enzima parcialmente disociado (BF_1 de *E.*

SUBUNIDADES (LIMITES PESO MOLECULAR) δ	ε	INHIBIDOR
(18.000-12.500)	(12.000-5.500)	10.500
(21.000-17.000)	(14.000-13.000)	
20.000	(13.000-11.000)	
20.000	12.000	
		¿25.000?
15.000	11.000	

des presentes y sus masas moleculares. En el caso de los factores F_1, CF_1 y BF_1 de *M. lysodeikticus* se han aislado las subunidades α y β, determinando las masas moleculares de los polipéptidos en cadenas al azar. La actividad de los factores BF_1 de *E. coli* y PS_3 se ha reconstituido a partir de las subunidades. El inhibidor natural de los factores F_1, CF_1 y BF_1 de *E. coli* se ha aislado.

coli) o de las subunidades aisladas, como se ha llevado a cabo con el factor F_1 de la bacteria termófila PS_3. Los resultados obtenidos muestran un cierto nivel de convergencia y permiten concluir una serie de generalizaciones a nivel de las subunidades mayoritarias. Los estudios realizados en diferentes laboratorios y con diferente enfoque experimental (Y. Kagawa y colaboradores; D. Deters, E. Racker y colaboradores, de la Universidad de Cornell; F. Mollinedo y V. Larraga en nuestro laboratorio) han permitido concluir que la subunidad β es esencial para la actividad de la ATPasa. Esta subunidad β requiere, sin embargo, una presencia mínima de subunidad α y una cierta conformación preexistente en la proteína activa. En efecto, tanto el grupo de Kagawa, estudiando el factor F_1 de la bacteria PS_3, como nuestro grupo, estudiando BF_1 de *M. lysodeikticus*, han demostrado que las subunidades α y β aisladas y los complejos de reasociación de las subunidades α y β son inactivos.

La identificación de la función de las subunidades minoritarias se encuadra en una situación menos definida y que, por el momento, permite escasas generalizaciones. En el caso del factor BF_1 de *E. coli*, se ha identificado que la subunidad δ es el polipéptido responsable de la reinserción del factor en la membrana. En ningún otro caso se ha podido establecer con un nivel de certeza semejante el papel de la subunidad δ. Conviene mencionar que un papel semejante ha sido postulado recientemente para este polipéptido en CF_1 (N. Nelson, Departamento de Biología, Technion-Israel Institute of Technology) o en la ATPasa de *St. faecalis*, donde el grupo de Abrams, de la Universidad de Colorado, había inicialmente demostrado que se requería un factor de peso molecular 37.000 dalton, llamado *nectina*, para la fijación de este BF_1 a la membrana. Para este último grupo, la *nectina* sería un componente funcionalmente análogo a la subunidad δ (20.000 dalton) en BF_1 de *St. faecalis*, constituyendo una forma dimérica de este polipéptido.

Muy recientemente, Abrams y colaboradores han puesto en evidencia que un pequeño fragmento de la subunidad α, liberado por digestión quimiotríptica de la ATPasa-F_1 de *St. faecalis*, es necesario para su inserción en la membrana. De la misma complejidad y falta de resolución en estos resultados conviene concluir precavidamente que una clara correlación entre subunidades de la ATPasa y su fijación a la membrana está aún por establecer. A este respecto, conviene mencionar que la situación es aún más complicada en el caso de F_1 mitocondrial donde se ha visto que otros polipéptidos, además de las 5-6 subunidades integrantes comunes del factor F_1, son necesarios para su reasociación con la membrana.

El papel de la subunidad γ es por el momento todavía más conflictivo. Basándose en métodos inmunológicos, Nelson, Racker y colaboradores han concluido que la subunidad γ de CF_1 está envuelta en la actividad fotofosforilante de la maquinaria transductora de energía de los cloroplastos. Los experimentos de reconstitución a partir de subunidades aisladas (caso del factor F_1 de la bacteria termófila PS_3) o de enzima parcialmente disociado (BF_1 de *E. coli*) han llevado a proponer que γ puede ser necesaria para un ensamblaje catalíticamente activo de las subunidades α y β. Es bastante sugestivo imaginar que el estudio de ATPasas de origen bacteriano cuyo holoenzima parece estar constituido por un número menor de subunidades puede ser decisivo para deducir el papel de las subunidades minoritarias. En este sentido, los trabajos realizados en nuestro laboratorio por F. Mollinedo y V. Larraga con BF_1 de *M. lysodeikticus* (constituido por subunidades α, β y γ), utilizando digestión tríptica controlada, indican que la subunidad γ de este F_1 puede desempeñar un papel esencial en su reasociación con la membrana.

A estas alturas se pueden emitir las siguientes conclusiones. En todos los factores F_1-ATPasas estudiados existe un núcleo central catalítico constituido por las subunidades mayoritarias α y β, siendo altamente probable que la subunidad β lleve el sitio catalítico propiamente dicho para la actividad ATPasa. La subunidad α puede contribuir a esta función conformacionalmente y/o por poseer los sitios de fijación para los nucleótidos. Si ello fuera cierto, la subunidad α intervendría de una manera decisiva en la actividad acoplante del factor F_1. El resto de las subunidades representan un conjunto mucho más variable y, por tanto, una propiedad menos estable de los factores F_1. Este conjunto de subunidades parece estar dirigido esencialmente a facilitar la interacción del núcleo central α β de cada F_1 con la membrana correspondiente, permitiendo que funcione eficazmente en el acoplamiento de energía y regulando su actividad hidrolítica para el ATP. Estas subunidades minoritarias deben formar parte de la puerta que regula el tráfico de protones, facilitado en la zona lipófila, impermeable, por el componente F_0. Por tanto, el conjunto de polipéptidos minoritarios debe ser un factor decisivo en la regulación de la función del complejo F_1 hacia la síntesis o la hidrólisis de ATP acoplando una u otra reacción con la entrada o la salida de H^+.

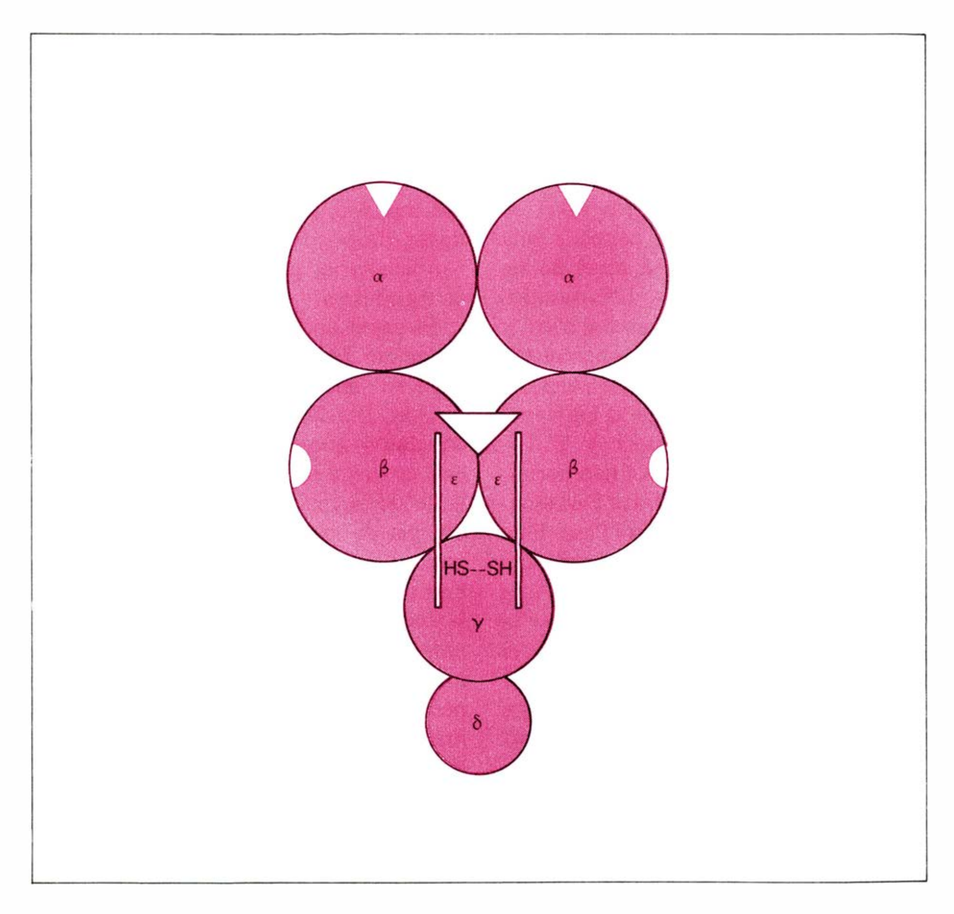
Hasta la fecha, es escaso el número de estudios que ha tratado de correlacionar subunidades con función acoplante, ya que se limita casi exclusivamente al realizado por Nelson, Racker y colaboradores. Uno de los problemas más inmediatos consiste en el diferente número de subunidades minoritarias que se identifican en factores F_1 de distintos orígenes. Dos son las hipótesis más razonables que se nos ocurren como posible explicación. Una, la más sencilla, es que todos los factores F_1 están formados por el mismo número de polipéptidos, pero cuya proporción y número varíe según el procedimiento de aislamiento o las manipulaciones a que se someten las proteínas. Otra hipótesis, en nuestra opinión más razonable, consiste en postular que diferentes tipos de subunidades pueden llevar a cabo una misma función, dependiendo de las características funcionales y estructurales del organismo u orgánulo de donde derive el factor F_1, es decir, como resultado de una profun-

da adaptación a las condiciones ambientales.

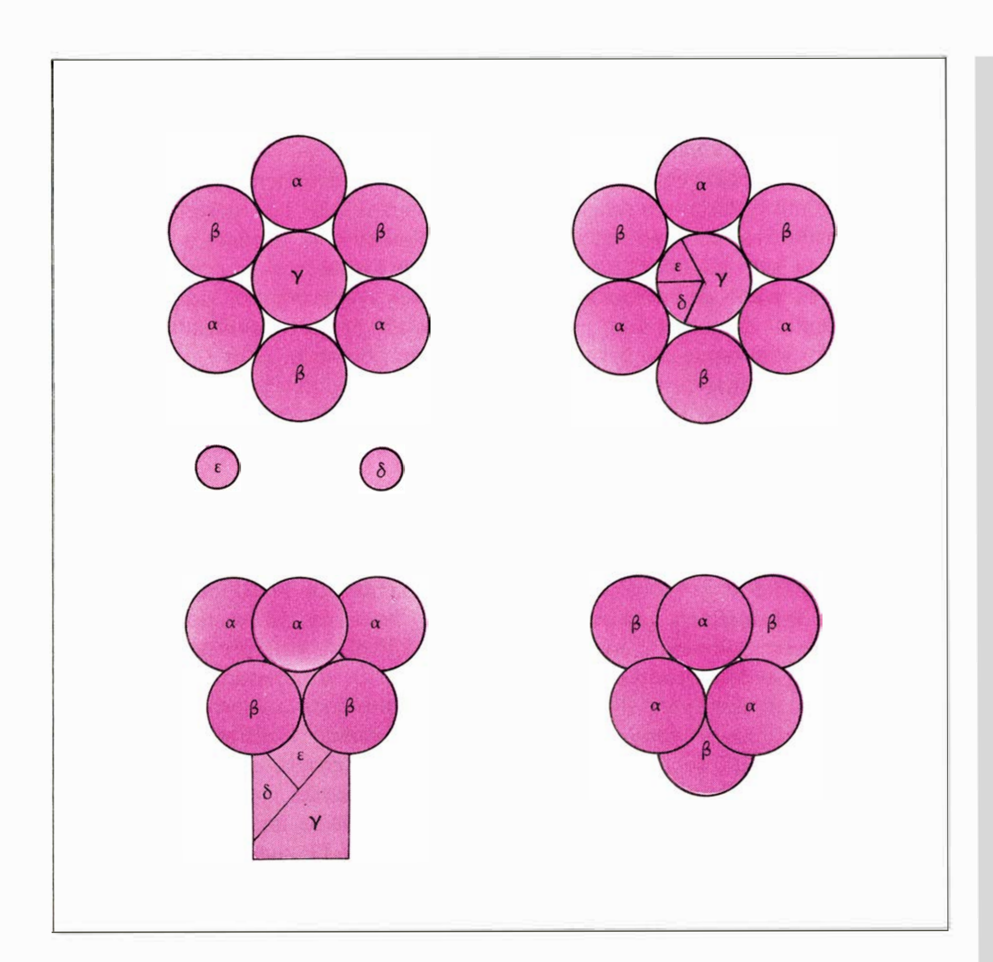
Para poder elegir entre estas hipótesis debe llevarse a cabo un análisis comparativo a nivel molecular entre diferentes factores F_1 adecuadamente seleccionados. La potencialidad de los factores F_1 purificados hasta el momento permite considerar la posibilidad con razonable optimismo. El esclarecimiento definitivo de la función de cada subunidad pasa por el estudio, tanto a nivel funcional como físico-químico, de las subunidades aisladas y sus interacciones. Este método es posiblemente el más lento, pero también el que puede ofrecer respuestas más racionales a las preguntas que hemos planteado.

Un paso ulterior en el avance de la estructura molecular de los factores F_1 consiste en la determinación de la estequiometría de sus subunidades, así como su distribución topográfica. Este tema es uno de los más conflictivos en la actualidad. Esta conflictividad contrasta con la línea de acuerdo a que se había llegado sobre este punto al principio de la década de los 70.

Criterios morfológicos, derivados del examen al microscopio electrónico de factores F_1 purificados, llevaron a pensar que una organización hexagonal con posible alternancia de las subunidades mayoritarias, con un elemento globular central formado por las subunidades minoritarias, constituiría el modelo más plausible. Esta idea se vio confirmada por otra vía experimental. La medida de la intensidad de coloración de las distintas subunidades, después de analizar los factores F_1 en geles con dodecil sulfato sódico, permitió aproximar la proporción relativa entre subunidades y con ello proponer fórmulas $\alpha_3\ \beta_3\ \gamma\ (\delta\ \varepsilon)$ para los factores F_1 de mitocondrias y bacterias (*M. lysodeikticus, E. coli* y la bacteria termófila fundamentalmente). En el caso de BF_1 de *E. coli* esta estequiometría ha sido confirmada por Bragg y colaboradores, de la Universidad de la Columbia Británica, Canadá, a partir de proteína purificada de células crecidas en presencia de una mezcla de aminoácidos marcados con ^{14}C, que habían sido añadidos al principio de la fase exponencial de crecimiento. Sin embargo, en el caso del factor CF_1 todas las pruebas disponibles hasta el momento señalan que la molécula de este F_1-ATPasa debe estar compuesta por 2α, 2β, 1γ, 1δ y 2ε. Recientemente se han postulado fórmulas $\alpha_2\ \beta_2\ \gamma_2\ \delta\ \varepsilon_2$ para el factor F_1 purificado de mitocondrias de corazón de buey y α_2



MODELO PARA EL FACTOR CF_1, propuesto por N. Nelson (del Technion Israel Institute), sobre la base de la estequiometría y función de sus subunidades ($\alpha_2\ \beta_2\ \gamma\ \delta\ \varepsilon_2$), recogiendo también las relaciones de proximidad entre las subunidades y la localización de dos grupos sulfhidrilo en la subunidad ε.



MODELOS PARA EL FACTOR F_1 MITOCONDRIAL. En la parte superior, aparecen dos modelos planos con alternancia de las subunidades mayoritarias y un núcleo central constituido por una copia de subunidad γ (P.L. Pedersen) o por una copia de γ + δ + ε (Kozlov y Mikkelsaar) que formarían el conjunto de una subunidad. La parte inferior de la ilustración muestra el modelo no plano con tallo propuesto por estos últimos autores (Kozlov y Mikkelsaar) en visión lateral y desde arriba. Para ellos, la transformación de una subunidad central en las tres subunidades γ, δ y ε resultaría de la acción de enzimas proteolíticos. En el modelo no plano, las subunidades α y β estarían en planos diferentes.

β_2 γ_2 δ_2 ε_2 para BF_1 de *E. coli.* A estas conclusiones se ha llegado por dos tipos de aproximación experimental. En un primer caso, por la localización y cuantificación de grupos sufhidrilo en las diferentes subunidades y, en el segundo caso, por la titulación de los complejos que intervenían en la reconstitución del enzima parcialmente disociado.

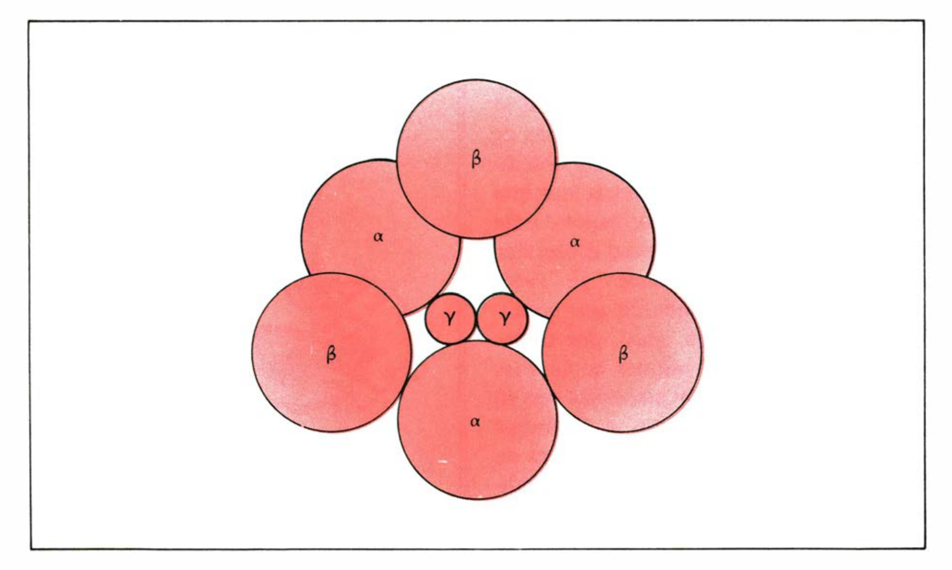
La reacción de las proteínas con agentes entrecruzantes constituye una poderosa técnica para analizar la organización geométrica de las subunidades de las proteínas oligoméricas, para determinar el número de subunidades que forman el oligómero y la relación especial entre subunidades, así como la posible simetría del oligómero. Su posible aplicación al estudio de la organización molecular de los factores F_1 encerraba una gran potencialidad que ha sido explorada en los últimos dos años. Aunque el análisis se ha realizado hasta el momento de modo excesivamente empírico e intuitivo ha rendido algunos resultados de interés. Por un lado, se ha confirmado la fórmula α_2 β_2 γ δ ε para el factor CF_1 con la que, si los pesos moleculares de las subunidades y del factor CF_1 en todo, son correctos, se obtiene una gran aproximación entre el peso molecular calculado (312.000 dalton) a partir de la estequiometría y el medido experimentalmente (325.000 dalton). Incluso queda espacio para alguna copia más de las subunidades más pequeñas (δ o ε). En el caso de F_1 de mitocondrias de corazón de buey, los resultados obtenidos apoyan una estequiometría mínima α_2 β_2 γ_2 δ ε. El cálculo correspondiente da para esta molécula un peso molecular de 321.500 dalton frente al valor 347.000 dalton determinado por equilibrio de sedimentación. Hay, por tanto, espacio para colocar alguna copia más de las subunidades más pequeñas con lo que podríamos acercarnos a fórmulas del tipos α_2 β_2 γ_2 δ_2 ε_2. La estructura de BF_1 de *Escherichia coli* se ha estudiado por métodos análogos. Bragg y colaboradores han llegado a la conclusión de la interacción entre las subunidades α - β, α - δ, β - γ, β - ε y γ - ε, pero no entre las subunidades más pequeñas ni entre

α - α y β - β. De acuerdo con un modelo plano, estas observaciones indican que las dos subunidades mayoritarias se distribuyen alternativamente. De acuerdo con un modelo en más de un plano, estos resultados pueden reflejar simplemente la distancia entre las subunidades. En cualquier caso, conviene tener en cuenta las limitaciones que encierran siempre los resultados negativos. Este comportamiento es semejante al del factor CF_1, pero difiere del de F_1. Este hecho es interesante porque la analogía funcional de la subunidad δ sólo se ha observado entre CF_1 y BF_1 de la bacteria *E. coli*, quedando por demostrar en otros factores F_1-ATPasas. En cualquier caso, no se ha llegado a proponer una disposición especial única para las subunidades de BF_1 de *E. coli*, manteniendo viva la controversia entre las dos fórmulas moleculares propuestas hasta el momento para este factor F_1 bacteriano.

Las diferencias observadas con esta aproximación experimental refuerzan la idea difundida por nosotros de que existen diferencias en la subestructura de distintos factores F_1, diferencias que puede ser muy útil explorar en detalle con objeto de correlacionar estructura, función y regulación de estos sistemas. Las grandes semejanzas en la composición de subunidades entre diferentes factores F_1 ha hecho presuponer un gran esfuerzo conservativo a través de la evolución y lógicamente suponer que la organización topográfica de las ATPasas debería haberse conservado igualmente. Esta conclusión nos parece un poco apriorística y debe considerarse con reservas, ya que se ha demostrado la existencia de diferencias en la organización molecular en proteínas funcionalmente idénticas, derivadas de diferentes orígenes como ocurre con la fosfofructoquinasa de levadura y la de mamíferos. Hay que plantearse además la posibilidad de que la diferencia en el número de subunidades minoritarias no refleje diferencias constitutivas en la subestructura de la proteína sino en diferencias accidentales en su grado de interacción con el núcleo central catalítico, que pueden depender, artificialmente, de la manipulación de las proteínas o, naturalmente, de un estado funcional. Estudios recientes en nuestro laboratorio sobre la biosíntesis del factor BF_1 de *Micrococcus lysodeikticus* (Carmen Muñoz, Universidad Complutense de Madrid) parecen apuntar hacia esta posibilidad.

Para terminar este punto conviene señalar que la información que se puede sacar sobre la estequiometría con el empleo de agentes entrecruzantes descansa en la necesidad de disponer de valores lo más exactos posibles del peso molecular de F_1 enteros y de sus subunidades. Como señalábamos anteriormente, esta cuestión es todavía debatible y exige que se lleve a cabo un esfuerzo experimental en esta dirección.

Por otro lado, existen limitaciones técnicas para abordar el problema de la organización y estequiometría de los complejos macromoleculares que son los factores F_1. Sin embargo, se abren perspectivas esperanzadoras con la aplicación de rayos X por dispersión a bajo ángulo y por estudios de difracción. Este último procedimiento requiere la cristalización de factores F_1, hecho que ya empieza a ser una realidad aunque no sabemos si será factible la obtención de derivados mesofórmicos. Otra posibilidad a explorar, utilizada recientemente por Kagawa, consiste en la aplicación de la difracción óptica a imágenes de la proteína al microscopio electrónico.



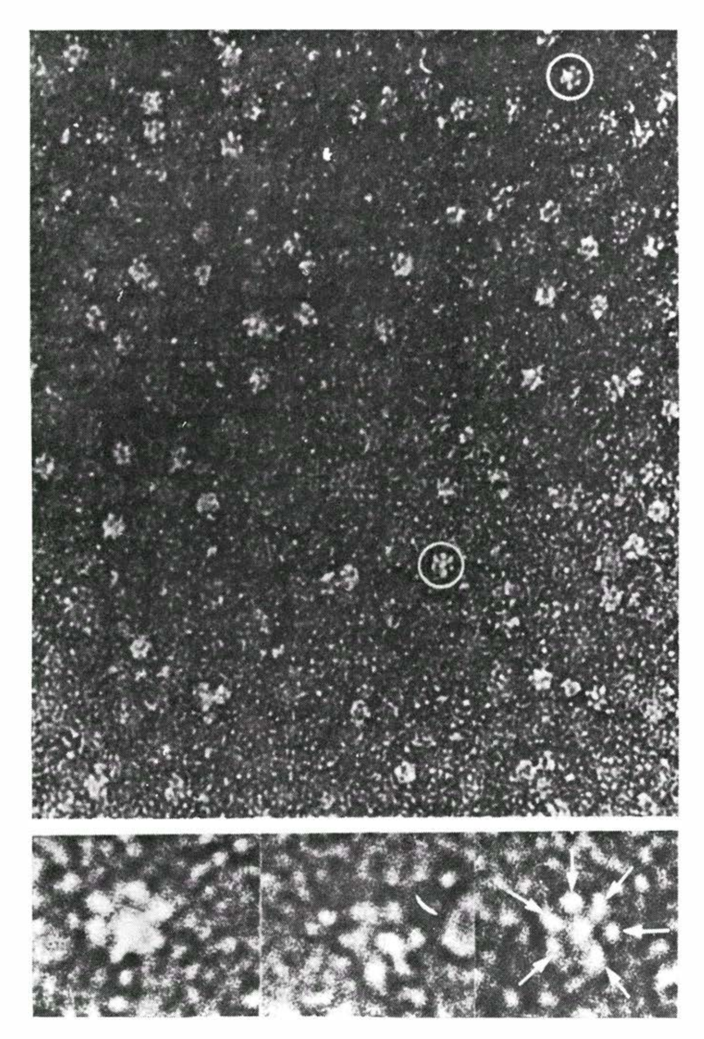
MODELO DEL FACTOR BF_1 de *Micrococcus lysodeikticus* mostrando una visión bajo un ángulo mayor que 90° desde la subunidad β. El modelo es asimétrico, con la distancia diferencial entre subunidades. Los círculos reflejan las masas moleculares de las subunidades y el número de copias.

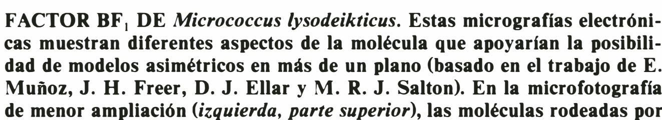
Muy recientemente se ha puesto de manifiesto la existencia de dos componentes no peptídicos que forman parte de los factores F_1 purificados. Uno de estos componentes, los nucleótidos, no están ligados covalentemente a la porción peptídica, pero sí con una alta afinidad, ya que copurifican con las ATPasas y sólo se liberan por desnaturalización y/o inactivación de la proteína.

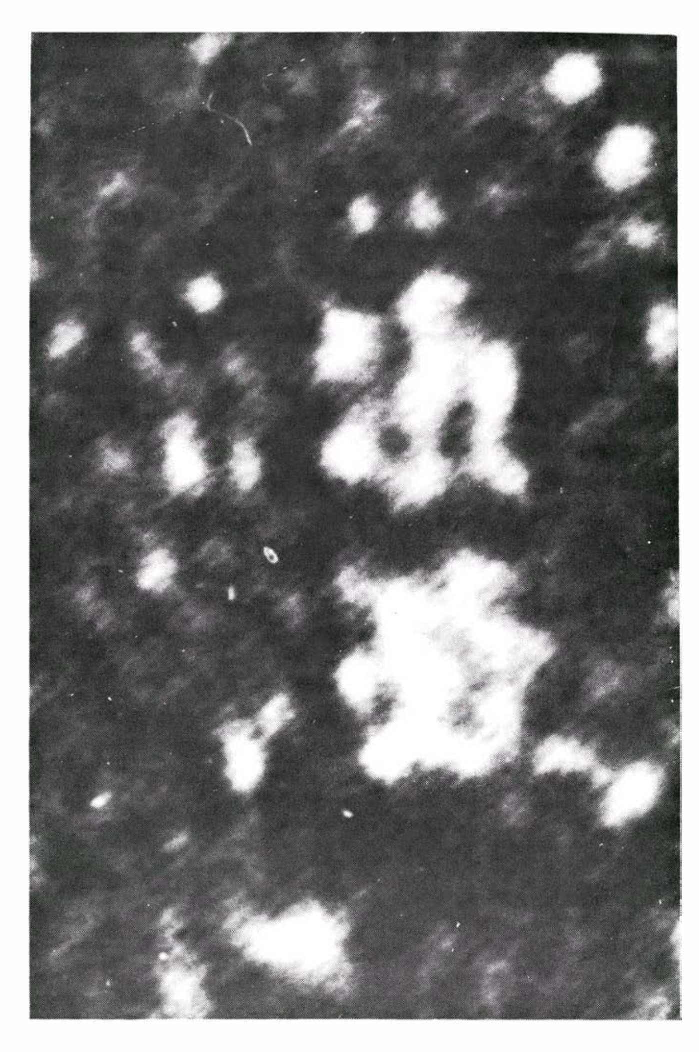
La presencia de estos nucleótidos firmemente asociados a las ATPasas hizo pensar que podían desempeñar un papel fundamental en la función de la proteína, tanto en su sentido hidrolítico como sintético. Este proceso ha sido investigado en el laboratorio de P. D. Boyer, examinándolo como posible apoyo de la hipótesis conformacional. El trabajo de este grupo, realizado sobre todo con mitocondrias y el factor F_1, ha demostrado que se forman cantidades catalíticas de ATP que permanece firmemente ligado a la proteína a partir de ADP y P.

En todos los factores F_1 -ATPasa estudiados, se ha detectado la presencia de nucleótidos íntimamente asociados, aunque el número de sus moléculas y su distribución entre ATP y ADP varía según el origen de la proteína. En F_1 se ha observado la presencia de estos dos nucleótidos en una relación de 3:2, lo que indica que esta ATPasa posee cinco sitios de alta afinidad. También se ha puesto de manifiesto que el número de nucleótidos fijados a la proteína aislada depende del método de separación y purificación de la proteína. De este modo, Garret y Penefsky, del Instituto para la Salud Pública de Nueva York, han conseguido obtener una preparación de F_1 sin ninguna molécula de nucleótido fijado, pero totalmente activa y capaz de fijar 5 moles de nucleótido.

En todos los otros casos se ha encontrado un menor número de moléculas de ATP y/o ADP por molécula de enzima: 1 mol de cada uno de ellos en BF_1 de *St. faecalis*, 1-2 moles de ATP y ADP en BF_1 de *E. coli*, 2 moles de ATP y 2 de ADP en BF_1 de *M. lysodeikticus*, mientras que CF_1 parece contener dos moles de ATP y 1 mol de ADP, aunque posee otro sitio de menor afinidad para este nucleótido. Dentro de la diversidad se observa que 4 es el número de sitios, dos para ATP y dos para ADP, en el que se encuentra mayor convergencia. Esta propiedad podría ser interesante en relación con el problema de la estequiome-







FACTOR BF_1 DE *Micrococcus lysodeikticus*. Estas micrografías electrónicas muestran diferentes aspectos de la molécula que apoyarían la posibilidad de modelos asimétricos en más de un plano (basado en el trabajo de E. Muñoz, J. H. Freer, D. J. Ellar y M. R. J. Salton). En la microfotografía de menor ampliación (*izquierda, parte superior*), las moléculas rodeadas por círculos presentan una apariencia hexagonal simétrica; pero en el resto del campo, ya se observa una menor simetría en las moléculas. Puede apreciarse un mínimo de siete glóbulos por molécula lo que apoya una estequiometría mínima α_3 β_3 $\gamma_{1\text{-}2}$, para BF_1 de *M. lysodeikticus*, suponiendo que cada glóbulo corresponde a una subunidad o a un conjunto de subunidades.

tría de subunidades que hemos discutido anteriormente, aunque nadie se ha planteado este problema hasta el momento. Naturalmente que para ello hay que determinar primeramente la capacidad máxima de fijación de nucleótidos y si las dos subunidades α y β poseen sitios de fijación independientes, que sean aditivos y no excluyentes. Por el momento, existen indicios de que esto es así, aunque no se ha demostrado inequívocamente. Conviene también advertir que la posibilidad de extrapolar el contenido de nucleótidos de un F_1 a la estequiometría de sus subunidades mayoritarias puede estar limitada por la falta de un equilibrio definido en el proceso de fijación de esos nucleótidos.

Otro aspecto interesante de las propiedades de los nucleótidos intrínsecos de los factores F_1 es su capacidad de intercambio con nucleótidos exógenos. Se ha comprobado que, en general, los nucleótidos intrínsecos intercambian difícilmente y que, en última instancia, este intercambio depende del estado energético de la proteína. Este hecho se ha puesto de manifiesto tanto en CF_1, en F_1 como en BF_1 de *E. coli* y *St. faecalis*. Sin embargo, BF_1 de *M. lysodeikticus* es capaz de intercambiar sus nucleótidos, lo que no está en contradicción con lo que antecede, ya que también se ha visto que en este BF_1 existe una relación entre nucleótidos fijados y actividad ATPasa. Este hecho podría relacionarse con el estado energético de la proteína, más elevado en unos factores F_1 que en otros.

Los hidratos de carbono representan la otra clase de constituyentes de algunos factores F_1, concretamente de CF_1 y BF_1 de *M. lysodeikticus*, como ha sido descubierto recientemente en nuestro laboratorio. El posible papel funcional de los azúcares en los F_1-ATPasas es menos obvio que el de los nucleótidos. Sin embargo, su papel estructural y su posible influencia en el ensamblaje de la proteína pueden ser decisivos. Una resolución a este problema impone una investigación destinada a demostrar la posible generalización de la naturaleza glucoproteica y a estudiar la topología de los azúcares en las moléculas de F_1. Como hemos hipotetizado recientemente, los azúcares pueden estar orientados de uno de los tres modos siguientes: (*a*) hacia el exterior de la proteína mirando hacia el citoplasma bacteriano, matriz mitocondrial o estroma de los cloroplastos; (*b*) en la misma localización con respecto a la proteína, pero dispuestos hacia la membrana facilitando así la interacción de los F_1 con la porción F_0; y (*c*) localizados en el interior de la proteína, formando un poro hidrofílico que podía ser factor esencial en el mecanismo de traslocación de protones.

El papel central que juegan los factores F_1 en el metabolismo energético, así como su complejidad funcional, permiten suponer que estos sistemas enzimáticos están sujetos a complejos procesos regulatorios. Se conoce poco con

respecto a la regulación de la síntesis de los polipéptidos constituyentes de F_1. Sin embargo, se ha descrito un buen número de variaciones posteriores a la transducción, algunas de las cuales modulan la función enzimática. Consisten en cambios en la estructura covalente, en ligandos unidos no covalentemente y en alteración de la conformación y de la estructura cuaternaria de estas proteínas. Por lo que se refiere a cambios en la estructura covalente, hemos observado que la glucoproteína ATPasa de *M. lysodeikticus* presenta, dependiendo del medio de crecimiento del microorganismo, formas cuyas partes sacarídicas tienen diferente composición de monosacáridos. La estructura covalente de los factores F_1 es también susceptible de cambios por efecto de degradación proteolítica o ruptura de enlaces lábiles (posiblemente enlaces glucosídicos) durante el curso de la purificación y posterior almacenamiento, si bien estas últimas modificaciones artificiales se pueden controlar experimentalmente. Existe, por tanto, la posibilidad de examinar diferentes estados de degradación producidos *in vivo*. Respecto a las diferencias en contenido de constituyentes no unidos covalentemente, hay que señalar que los factores F_1 presentan una ocupación variable de los sitios de fijación de nucleótidos, que se pueden correlacionar con niveles de actividad enzimática, como se ha expuesto en el apartado anterior y que depende, para un determinado material, del estado energético de la membrana de partida y del procedimiento de purificación del factor F_1.

La molécula del factor de acoplamiento en solución es susceptible de cambios conformacionales, inducidos por unión de ligandos (ADP, ATP, el antibiótico aurovertina, el alcaloide quercetina, por cationes divalentes, por oxianiones y por cambios de *p*H). Estos cambios se han puesto de manifiesto bien por cambios en las características cinéticas o por existencia de transiciones lentas en la reacción de hidrólisis de ATP y en la unión de ligandos, como se ha comprobado al estudiar la accesibilidad de F_1 a enzimas proteolíticos, sus propiedades ópticas como dicroísmo circular y dispersión rotatoria, directamente relacionadas con la estructura secundaria de la molécula. L. Cantley y G. Hammes, de la Universidad de Cornell, han propuesto la regulación alostérica de los sitios de hidrólisis de ATP por sitios de fijación no hidrolítica de nucleótidos. Toda esta flexibilidad conformacional interviene probablemente *in vivo*, ya que el factor de acoplamiento CF_1 de cloroplastos asociado a membrana muestra cambios conformacionales (intercambio de hidrógeno y exposición de grupos funcionales) dependiente del estado energético de la membrana (formación de gradiente de protones por iluminación) como han observado I. Ryrie, A. Jagendorf y McCarty, de la Universidad de Cornell.

Las variaciones en la estructura cuaternaria de factores de acoplamiento F_1, centradas fundamentalmente en la presencia o no de cada subunidad minoritaria, pueden atribuirse a los diferentes organismos de partida y pérdidas durante los procedimientos de purificación. Sin embargo, el mismo microorganismo *M. lysodeikticus* es capaz de producir diferentes formas moleculares de BF_1 que, una vez purificadas de acuerdo con el mismo procedimiento, poseen o carecen del polipéptido inhibidor de la actividad ATPasa, como ha puesto en evidencia J. Carreira en nuestro laboratorio. Esto no indica la carencia *in vivo* de esta subunidad asociada con la ATPasa, pero sí una interacción de diferentes características entre el inhibidor y la ATPasa.

Algunas de las modificaciones que hemos comentado (exceptuando las modificaciones en estructura cuaternaria) son las causas que probablemente conducen a que, en preparaciones de factores de acoplamiento purificado y homogéneo de acuerdo con criterios convencionales, coexistan subespecies moleculares que difieren ligeramente en su tamaño hidrodinámico efectivo y/o carga (separación por electroforesis de alta resolución), como hemos observado en nuestro laboratorio con BF_1 de *M. lysodeikticus*, e independientemente E. Moudrianakis, de la Universidad Johns Hopkins, con BF_1 de *Alcaligenes faecalis*, CF_1 de cloroplastos y F_1 mitocondrial.

Como definíamos al principio de este trabajo, el factor F_0 posee un marcado carácter hidrofóbico, lo que ha dificultado su caracterización molecular y su análisis bioquímico. Sin embargo, su importancia es obvia en la vía de traslocación de protones y en este sentido se está potenciando el trabajo sobre este factor. Hasta este momento, las preparaciones F_0 de mitocondrias y la de la bacteria termófila PS_3 son las mejor caracterizadas y con las que se ha podido demostrar que forman una vía específica conductora de protones que se bloquea por el factor F_1 o, artificialmente, por el reactivo diciclohexilcarbodiimida. Este reactivo actúa como inhibidor relativamente específico de la actividad ATPasa de los factores F_1 cuando están unidos al componente F_0. Por tanto, se ha postulado que el sitio de sensibilidad para este reactivo está localizado en el componente F_0. La inhibición es irreversible y se ha presupuesto que es covalente. Sacando partido de este hecho, se ha logrado purificar la proteína que reacciona con la carbodiimida de BF_0 de *E. coli* y sugerir con bastante fuerza que la función del factor F_0 es actuar como una traslocasa de protones. Esta proteína se ha purificado después en estado nativo por el grupo de Gitler y N. Nelson en Israel a partir del complejo CF_1-CF_0, y se ha caracterizado como un proteolípido que forma un canal a través del cual se transportan los H^+ en las membranas transductoras de energía.

La transducción de energía, uno de los problemas fundamentales en biología y que ha desafiado, durante más de 20 años, los progresos de la bioquímica, ha encontrado al parecer el camino hacia una solución. Se ha llegado a una situación de acuerdo para abordar el problema a nivel molecular, tratando de profundizar en el conocimiento de la relación estructura-función de los catalizadores que intervienen en el proceso. Ello ha supuesto superar una barrera experimental que durante muchos años se ha consagrado a examinar la cuestión sólo fenomenológicamente.

Como hemos visto antes, los complejos F_0-F_1, ATPasas constituyen sistemas altamente ordenados en los que la estructura, conservada a través de la evolución, se ha adaptado a las peculiaridades fisiológicas para jugar con la máxima eficacia su papel en una de las funciones esenciales para el desarrollo de los seres vivos: la transformación de la energía en formas asimilables. Como prueba del carácter conservativo de estas estructuras hemos definido una serie de propiedades comunes a todos los factores F_1 estudiados hasta el momento y procedentes de mitocondrias, orgánulos fotosintéticos y membranas plasmáticas de bacterias. La información sobre los componentes F_0 es menos completa por el momento, por lo que permite menor número de generalizaciones. Sin embargo, estos estudios han puesto también de manifiesto la existencia de diferencias, probable consecuencia de la divergencia funcional. La exploración de estas diferencias utilizando sistemas de mutantes y diferentes cepas de microorganismo puede ser el factor decisivo en el esclarecimiento de las bases moleculares de la transducción de energía, un proceso esencial para el desarrollo y la conservación de la vida en nuestro universo.

Juegos matemáticos

¿Se trata de un robot superdotado, o del doctor Matrix, que vuelve a la carga?

Martin Gardner

"En realidad, los ordenadores no piensan. Lo que ocurre es que se piensa que piensan. (Pensamos nosotros.)"
—Theodor H. Nelson, *Dream Machines*

Durante años, científicos expertos en ciencias de computación y en lingüística se han esforzado en poner a punto sistemas que permitan a los ordenadores electrónicos aceptar entradas y emitir salidas en lenguaje hablado ordinario. Los avances conseguidos en esta dirección han sido, sin embargo, descorazonadoramente lentos. Hoy por hoy solamente pueden desarrollarse los más triviales diálogos, por lo general mediante frases estereotipadas de léxico extraordinariamente reducido, mecanografiadas en el ordenador.

Por otra parte, hace más de medio siglo que los aficionados a la ciencia ficción están habituados a los ordenadores parlantes. Estas maravillosas máquinas han aparecido incluso en la literatura fantástica infantil. Ya en 1907 Frank L. Baum introdujo (en *Ozma of Oz*) un robot mecánico de cuerda llamado Tik-Tok, que según sus fabricantes podía pensar, hablar, actuar y "hacer de todo, menos vivir". Más recientemente, a través de filmes como *El planeta prohibido (Forbidden Planet)* y *La guerra de las galaxias (Star Wars),* el público se ha familiarizado con los robots y computadores parlantes, como HAL, que gobernaba la nave espacial de *2001, una odisea del espacio (2001: A Space Odissey)*, escrita por Arthur C. Clarke.

Hay numerosos signos demostrativos de la naturalidad con que ha sido aceptada esta idea. Tomemos por ejemplo un robot de juguete que se puso a la venta el año pasado: el niño responde a ciertas preguntas de elección múltiple pulsando botones en el dorso del robot, y el robot comenta (mediante una cinta magnética previamente grabada al efecto) la calidad de las respuestas. Además, los programas de ordenador capaces de jugar partidas de ciertos juegos de salón están mejorando muy rápidamente. A un coste relativamente barato puede adquirirse un computador ajedrecístico capaz de ganarle a un jugador novel. Los programas de ajedrez más perfectos alcanzan ya el nivel de maestro; de hecho, acortando suficientemente el tiempo disponible entre jugadas, tales programas pueden propinarle una zurra incluso a un gran maestro.

Debido a estas y otras razones, no hubo excesivo escepticismo cuando el pasado mes de julio se anunciaron en Estados Unidos las primeras exhibiciones públicas de ASMOF, el primer ordenador parlante del mundo. El nombre ASMOF era acrónimo de la American Superior Mind Operating Foundation, que patrocinaba las exhibiciones del robot prototipo para estimular las ventas de sus productos. Me enteré por recortes de prensa que me fueron enviando los lectores de que el nuevo robot disponía de una nueva memoria de burbujas magnéticas. Sus circuitos estaban alojados en un cuerpo de aluminio de 6 metros y medio de alto, cuyo diseño recordaba el de los robots cinematográficos. El monstruo tenía dos aberturas en los lugares correspondientes a los ojos de un ser humano, y un tercer ojo provisto de una lente color rubí en el medio de su enorme frente. No tenía nariz, aunque un altavoz cónico formaba a modo de boca. En las exhibiciones ante el público el robot estaba sentado tras una enorme mesa, y permanecía inmóvil. De cuando en cuando lanzaba un rayo de luz granate para explorar la mesa o para observar a alguien situado frente a él.

ASMOF comenzó su gira por el país a finales de julio, haciendo sesiones de dos horas en salas y teatros de grandes ciudades y enclaves turísticos. Pagando 3 dólares de entrada los espectadores tenían derecho a sentarse frente a ASMOF y formularle una pregunta razonablemente breve sobre cualquier tema. El robot respondía con voz metálica y gangosa. A veces, antes de responder, el robot se paraba a bromear con el interrogador. No lo sabía todo: en ocasiones respondía con observaciones del tipo "Su problema requería un tiempo de cálculo excesivo" o "Lo lamento, señora, pero tal información no consta en mi banco de memoria".

En ocasiones especiales, el robot dedicaba una sesión completa a combatir contra algún experto local en algún juego intelectual de mesa. ASMOF parecía ser jugador de primera categoría en ajedrez, damas, go, e incluso en juegos con elementos aleatorios, como el backgammon, el bridge y el póker. A principios de agosto, cuando el robot se encontraba en un teatro de Nyack, Nueva York, le fueron ofrecidos 1000 dólares a un gran maestro de ajedrez de los alrededores si conseguía ganar a ASMOF. El gran maestro perdió en 18 jugadas. ASMOF retó entonces al jugador a un segundo encuentro, ofreciendo jugar sin el alfil de dama si el gran maestro se avenía a apostar 1000 dólares en el juego. Tal oferta fue declinada.

La ignominiosa derrota del gran maestro dio a los acontecimientos un giro sorprendente. Sin embargo, cuando telefoneé a un viejo amigo que trabaja en el laboratorio de inteligencia artificial del Instituto de Tecnología de Massachusetts (MIT), éste me aseguró que ASMOF era un timo completo. No tenía certeza de cómo el robot era controlado, pero sí estaba seguro de que la fundación que lo respaldaba era una tapadera. Confidencialmente sugirió que quizá mi antiguo conocido, el doctor Joshua Irving Matrix, pudiera estar detrás de todo el asunto.

Aunque ojeé mi colección de recortes de prensa no pude encontrar ninguna fotografía del director de la fundación, Frank Rossum, ni de su ayudante, Josie Clarke Nelson. A pesar de ello, y conociendo la obsesión del doctor Matrix por los juegos de palabras, empecé a reflexionar sobre ambos nombres. ¿Podría ser Frank abreviatura de Frankenstein? Desde luego, Rossum hace pensar en seguida en los Robots Universales de Rossum, de la obra teatral *R.U.R.,* escrita en 1920 por Karel Čapek. (Esta obra introdujo en el idioma inglés el término "robot", procedente del checo *robota*, que significa trabajo forzado, servicio obligatorio.) ¿Qué pensar de Josie Clarke Nelson? Josie podría ser una versión femenina de Joe, el famoso robot de las obras de ciencia-ficción de Lewis Padgett. Clarke, evi-

dentemente, no podía referirse sino a Arthur C. Clarke. Y Nelson podría guardar relación con Theodor H. Nelson, joven especialista de ciencias de computación cuyo libro doble, *Dream Machines* y *Computer Lib* está conociendo un éxito colosal, siendo como es una introducción –que recuerda a los libros de *Oz*– al fantástico nuevo mundo que los ordenadores están creando. (Su último libro, *The Home Computer Revolution,* contiene predicciones todavía más asombrosas.)

¡Pero aguarde un momento! Había todavía otra extraordinaria coincidencia. Fue el propio doctor Matrix quien reveló que HAL se deduce de IBM sustituyendo cada letra por su precedente. Si en vez de hacerlo así tomásemos la siguiente ¡se obtendría *JCN*, anagrama del nombre de la ayudante de Rossum!

Por un folleto facilitado por la propia fundación Rossum pude saber que la próxima actuación de ASMOF tendría lugar en el anfiteatro de la venerable Chautauqua Institution, en la región norteña del estado de Nueva York. Un día de carretera me bastó para llegar desde mi casa de Westchester hasta un motel cercano a la entrada de la Chautauqua Institution. Por la tarde del día siguiente puse el mayor interés en llegar al anfiteatro lo bastante pronto como para asegurarme una localidad de primera fila. Previendo la posibilidad de que el doctor Matrix o su preciosa hija medio japonesa pudieran presentarse en el estrado me disfracé con unas gafas obscuras y gran mostacho postizo.

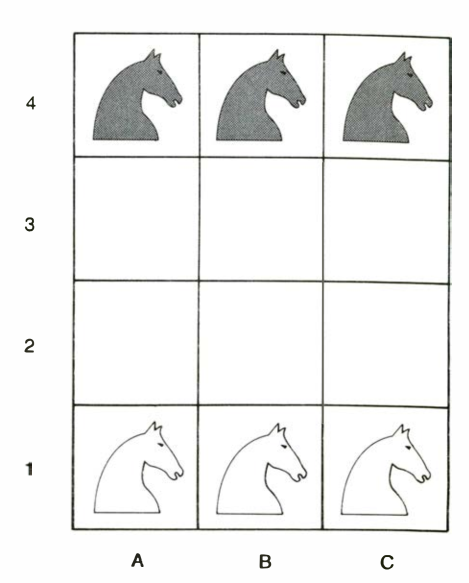
No dieron señal de vida. Un joven de agradable aspecto, largos cabellos pelirrojos y barba rojiza se presentó a sí mismo como "futurólogo" y ex-colaborador del Stanford Research Institute, y nos expuso doctamente los recientes progresos alcanzados en la comunicación con ordenadores mediante lenguajes naturales, para terminar examinando la monumental brecha abierta por los investigadores en ciencias de computación al servicio de la American Superior Mind Operating Foundation. Los decubrimientos de estos investigadores, hechos posibles por las nuevas memorias de burbuja magnética, permanecían todavía clasificados como alto secreto. De forma velada, dejó entrever que la CIA estaba intentando impedir la manufactura de versiones reducidas de ASMOF. Nos aseguró que a diferencia de los humanos, los gatos y las cucarachas, el robot no poseía percepción extrasensorial. Por consiguiente, ASMOF no respondería ninguna pregunta sobre el futuro ni tampoco aquellas que pudieran requerir algún tipo de PES. La fundación trabajaba en estrecho contacto con varios físicos del Stanford Institute, al objeto de incluir en la circuitería de ASMOF elementos de precognición y percepción extrasensorial, aunque, añadió, tales perfeccionamientos podrían requerir otra década.

Fue una sesión original y entretenida. Las preguntas de la mayoría de los asistentes fueron muy triviales e ingenuas, como por ejemplo, durante qué años fue Millard Fillmore décimotercer Presidente de los Estados Unidos, o qué equipo ganó la eliminatoria final de béisbol en tal y cual temporada, datos todos ellos fácilmente almacenables en la memoria de un ordenador. Algunas preguntas, no obstante, fueron más difíciles. Por mi parte, tomé buena nota sobre las relativas a juegos de palabras y matemática recreativa, temas que podrían interesar a la mayoría de los lectores de esta sección.

Por ejemplo, una jovencita le preguntó a ASMOF cuál era la palabra más larga del idioma español. Tras escrutar a la jovencita con su ojo color rubí y alabar su simpática sonrisa, ASMOF preguntó: "¿Incluidas las palabras compuestas?" Al contestar la muchacha que sí, ASMOF replicó: "En tal caso, no existen palabras de longitud máxima. Siempre podría hablarse de tatarabuela, tatara-tatarabuela, y así sucesivamente. La siguiente pregunta, por favor".

He aquí algunas otras curiosidades lingüísticas, no muy conocidas. Una señora quiso saber si existía alguna palabra monosilábica (inglesa) de 10 letras. ASMOF consiguió salir del paso con "scraunched". Un joven le preguntó si existía alguna palabra inglesa que contuviera en sucesión las letras "nkst". Tras algo de cháchara sobre la leyenda que el muchacho llevaba impresa en la camiseta, ASMOF dio una respuesta satisfactoria. A continuación intervino su hermano, quien formuló una pregunta análoga, esta vez con "nksh". ¿Podría el lector encontrar en un diccionario sendas palabras que contengan estas sucesiones, antes de que el mes próximo aparezcan publicadas las soluciones en esta sección?

A los niños les gustaba proponer al robot acertijos absurdos. A veces ASMOF conseguía dar respuestas correctas. Cuando no ocurría así, el robot preguntaba siempre la solución, con objeto de almacenarla en sus circuitos de memoria, alabando en ocasiones al niño que había conseguido dejarle sin respuesta. Entre mis notas encuentro que un señor mayor, profesor emérito de inglés en la cercana Universidad del Estado de Nueva York en Buffalo, preguntó al robot un acertijo literario que me era totalmente desconocido. Quiso saber qué relación tenían el viejo marino de Coleridge y un jugador de béisbol borracho. ASMOF respondió rápidamente: "He stoppeth one of three" (paraba una de cada tres).



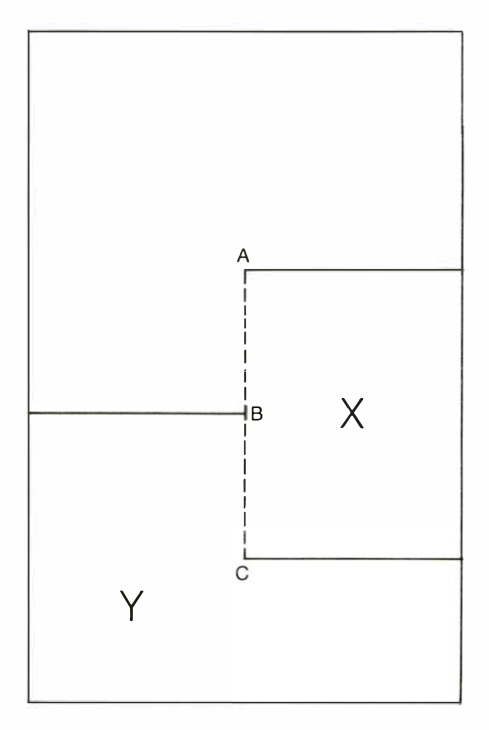
Problema de la permutación de caballos

La pregunta más difícil sobre juegos de palabras se la formuló una especialista en Shakespeare de la facultad de Bryn Mawr. En una de las obras de Shakespeare, afirmó, hay un verso que comienza "My". Leyendo hacia arriba, las primeras letras de los cuatro versos anteriores dan la palabra WANT. Leyendo hacia abajo, las primeras letras de los cuatro versos sucesivos dan BABY, con lo que resulta, en definitiva, el acróstico WANT MY BABY. ¿En qué obra se encuentra este acróstico?

La pregunta debió ser muy difícil, pues ASMOF dedicó varios minutos a discursear sobre la intencionalidad o casualidad de tales acrósticos en la obra de Shakespeare. Finalmente, el robot acertó a localizar el verso. Se trata de la decimocuarta línea, contada desde el final de la primera escena del primer acto de *The Comedy of Errors*: "My soul should sue as advocate for thee". Pudiera ser, conjeturó ASMOF, que Shakespeare estuviera enviando una súplica en clave a alguna dama amiga, buscando protección para algún hijo ilegítimo.

Algunas preguntas tenían relación con la lógica formal. Un señor de New Jersey llamado Ken Knowlton intentó hacer que el robot entrase en un círculo vicioso de respuestas alternativamente afirmativas y negativas, proponiéndole una conocida paradoja de Bertrand Russell, a saber, la del barbero de pueblo, que afeita a todos los hombres y solamente a aquellos hombres que no se afeitan a sí mismos. "¿Se afeita el barbero a sí mismo?", preguntó Knowlton. ASMOF replicó: "Insuficiencia de datos. El barbero podría ser mujer".

¿Cuál es el único procedimiento –quiso saber un matemático de Princeton– para clasificar todos los enteros positivos en dos conjuntos disjuntos, de manera que ningún par de números de un mismo con-



Cómo doblar la hipercarta

junto tenga como suma un número primo? Daré la sencilla solución de este problema en el número del mes que viene. Por cierto, que el robot no supo hallarla.

Thomas Szirtes, de Montreal, quiso saber de ASMOF si era posible que él (Szirtes) tuviese exactamente un tercio de escocés, un tercio de chino y un tercio de húngaro. ASMOF explicó que no. Cada persona, dijo el robot, tiene 2^1 padres, 2^2 abuelos, 2^3 bisabuelos, y así sucesivamente. Por consiguiente, la pregunta equivale a saber si 2^n puede ser igual a $3x$. Ahora bien, 2^n es igual a $2.2.2....2$, con el 2 repetido n veces. Como 2 es número primo, sabemos por el teorema fundamental de la aritmética que esta sucesión es la única descomposición de 2^n en producto de factores primos, por lo que 2^n no tiene ningún divisor primo distinto de 2. Así pues, 3 no puede ser divisor de 2^n, y la conjetura original tiene que ser falsa.

Jaime Poniachik, de Buenos Aires, editor de una excelente revista de rompecabezas matemáticos titulada *The Snark* se encontraba casualmente visitando la Chautauqua Institution cuando se celebró la exhibición de ASMOF. Llegado su turno de preguntar, Poniachik dijo tener en los Estados Unidos un amigo con un curiosísimo número de Seguridad Social: en sus nueve cifras se encuentran todos los dígitos de 1 a 9, formando un número cuyas dos primeras cifras (leídas de izquierda a derecha) dan un número divisible por 2; las tres primeras cifras componen un número divisible por 3; las cuatro primeras, uno divisible por 4; y así sucesivamente, siendo el número completo, de nueve cifras, divisible por 9. ¿De qué número se trata? ASMOF manifestó su admiración por este nuevo problema, alegando falta de tiempo para determinar la solución. Cuando volvía a casa, yo mismo encontré la única solución, pero la reservaré para el mes próximo.

Llegado mi turno, pedí permiso para proponer un problema donde intervienen seis caballos de ajedrez, en un tablero de tres por cuatro. En una hoja de papel esbocé la posición inicial, que se muestra en la ilustración de la página anterior. El problema consiste en permutar entre sí las posiciones de los caballos blancos y negros empleando el número mínimo de movimientos. Los caballos pueden moverse en cualquier orden, sin que importe su color, pero no se permite que haya dos ocupando simultáneamente la misma casilla.

Cuando este problema fue publicado por primera vez en 1974, en el *Journal of Recreational Mathematics*, fue declarado trivial, dándose allí una solución en 26 movimientos. Más tarde se descubrió una solución en 18 jugadas. Este problema figura en la primera edición de mi último libro, *Aha!* (manual de una colección de seis filminas para enseñanza secundaria, llamada *The Aha! Box*) y en él se menciona que la solución óptima consta de 18 jugadas. Tres lectores del libro, Gary Goodman, Warren B. Porter y George Schneller consiguieron, independientemente, reducir la solución a 16 movimientos.

Tras escrutar mi boceto con su tercer ojo, ASMOF presentó la solución de 18 jugadas, poniendo en duda mi afirmación de que había otra más breve. Tras exponerle la solución abreviada, ASMOF felicitó a sus descubridores por su perspicacia, al darse cuenta de cómo evitar un movimiento en apariencia redundante. ¿Podrían los lectores descubrir la solución en 16 pasos, antes de que sea revelada el mes que viene?

Al cabo de más o menos una hora de exhibición, el espectáculo me había convencido de que mi amigo del MIT estaba en lo cierto. Ninguno de los cálculos se estaba realizando en los circuitos internos del robot; con seguridad el robot era controlado desde algún lugar idóneo cercano. Pude enterarme, por un equipo de informadores de la cadena de televisión ABC allí destacado para grabar el acontecimiento, de que Rossum y Nelson se habían alojado en un hotel cercano a la Chautauqua Institution. Sin llamar la atención me escurrí fuera del anfiteatro. Cuando llegué al hotel de que me hablaron, el vestíbulo y los pasillos se encontraban desiertos. Al parecer, todo el mundo se había ido a la exhibición. Nadie me vio aplicar la oreja a la puerta de la habitación de Rossum. Se podían percibir, amortiguadas, las voces de una conversación.

El hotel ya no es nuevo, así que tuve pocas dificultades para introducir por la rendija de la puerta una tarjeta de crédito de plástico y forzar hacia atrás el pestillo. La escena que se desarrollaba en la habitación no era demasiado diferente de lo que había previsto. Iva estaba sentada frente a un gran pupitre, y llevaba puestos unos enormes auriculares. Un aparato de televisión en color situado frente a ella mostraba al espectador que en ese momento interrogaba a ASMOF. Más tarde me enteré de que la cámara de televisión estaba disimulada tras el ojo izquierdo del robot. Ambos lados del pupitre de Iva estaban ocupados por los libros de referencia habituales, que los bibliotecarios suelen tener a mano para responder a consultas telefónicas: *El Almanaque Mundial, La Enciclopedia Espasa*, varios *Quién es Quién*, un diccionario completo, concordancias de la Biblia y de autores clásicos, diccionarios de citas, etc. También tenía al alcance de la mano la consola de un gran ordenador programable y un teléfono especial, donde para marcar un número bastaba introducir una tarjeta perforada.

Iva se sobresaltó al entrar yo. Se quitó los auriculares y tapó con la mano el micrófono que tenía frente a sí. "Si no se va inmediatamente, gritaré y llamaré a la policía", dijo airadamente, con los ojos centelleantes.

"Me parece que no", repliqué, quitándome las gafas obscuras y arrancándome el bigote postizo.

"¡Otra vez tú!", dijo Iva haciendo una mueca. "Siéntate por ahí y mantén cerrada la bocaza."

Me senté en el borde de la cama, y observé su trabajo durante el resto de la sesión. Un dispositivo electrónico deformaba completamente la agradable voz de Iva, dándole el timbre frío y mate característico de los robots de cine y televisión. Las preguntas fáciles presentadas por almas simples obtenían rápida respuesta, pues ella albergaba en su cabeza una nada despreciable cantidad de información. Otras preguntas requerían un rápido vistazo a algunos de los libros de consulta del pupitre, mientras para ganar tiempo Iva se enzarzaba de cháchara con el espectador de turno. Para las preguntas que exigían cálculos y operaciones aritméticas, como hallar la decimotercera raíz de 10, o elevar el número "pi" a la potencia "pi", y otros semejantes, disponía del computador. Cuando la pregunta era más difícil recurría al teléfono, insertando en él una ficha. Al volver a enunciar el problema en voz alta —ostensiblemente para cercio-

rarse de que lo había comprendido correctamente– lo que hacía Iva era preguntárselo a su interlocutor telefónico. Más tarde, Iva me dijo que el doctor Matrix había contratado a más de 100 especialistas, a los que pagaba espléndidamente, cada uno de ellos experto en un tema, para permanecer en contacto telefónico a través de números particulares durante las sesiones de exhibición.

Iva me dijo también que el doctor Matrix había intentado contratar a Isaac Asimov para que hiciera el trabajo que ella realizaba, evitando así tener que contratar tantos expertos. Sin embargo, Asimov había declinado cortésmente la proposición, en parte, debido a que tenía que terminar 17 libros antes de fin de año. Nunca conseguí enterarme de cómo se llamaba su especialista en matemáticas; tampoco pude sonsacarles quién fue el gran maestro de ajedrez que jugó las partidas.

Terminada la exhibición, Iva se quitó los auriculares, apagó el micrófono y sonrió. Miró furtivamente su reloj de pulsera, uno de esos relojitos miniatura que son tan caros que es imposible consultar las agujas sin ayuda de una lente de aumento. "Es hora de tomar una copa", dijo.

Estuve encantado de saber que el doctor Matrix había tenido que salir de la ciudad aquel día. El viejo buitre había ido a visitar a un amigo que se ganaba la vida haciéndose pasar por "medium" en la vecina comunidad espiritista de Lily Dale. Como a los restaurantes situados en torno a la Chautauqua Institution no se les permite servir bebidas alcohólicas (residuo del piadoso pasado de la Institución), Iva me condujo hacia el norte por la Nacional 17 hasta un pintoresco restaurante de mariscos, en Westfield, a orillas del Lago Erie. (Me llamó la atención el siguiente reclamo de su carta de platos: "Servimos ostras en Enero, Febrero, Marzo, Abril, Maryo, Jurnio, Jurlio, Argosto, Septiembre, Octubre, Noviembre y Diciembre".) Encargamos la cena.

"En mi opinión", dije entre dos bocados, "no es más que la ensalada de un mal café".

Iva dio un respingo. "Pues en mi opinión, la ensalada es excelente", dijo. A pesar de todo tuvimos una cena deliciosa, y pasamos juntos una velada muy agradable.

Un mes más tarde, estando ASMOF contratado para actuar en un teatro de Washington, dos periodistas recelosos, ayudados por Randi el Mago, descubrieron toda la superchería. Alquilaron una habitación contigua a la que Iva utilizaba para manejar el robot, taladraron un agujero en la pared divisoria, y consiguieron buenas fotografías de Iva sentada al micrófono. La historia apareció publicada en la edición matinal del diario, pero para entonces el doctor Matrix e Iva habían ya desaparecido. ASMOF quedó abandonado en el camión que utilizaban para transportarlo. La policía de Washington ha entregado el robot a la Smithsonian Institution donde pronto será exhibido con una reseña magnetofónica de todo el timo.

La "hipercarta", que presenté el mes pasado como pieza de escultura minimal, se construye fácilmente como sigue: se toma una pieza rectangular de cartulina o papel fuerte, y se practican tres cortes a lo largo de las tres líneas negras indicadas en la ilustración de la página anterior. A continuación se dobla la solapa X 90 grados a lo largo de la línea de puntos ABC y después se vuelve del revés la porción inferior Y doblándola hacia atrás 180 grados a lo largo de la línea de puntos BC.

Ya mencioné en el pasado mes de mayo que Alan Cassel había descubierto, mediante un método "de Monte Carlo" para determinar si un número dado es primo, que las probabilidades de que el número formado por 123456789 repetido siete veces y seguido por 1234567 fuese primo eran de alrededor de un trillón a uno. Un físico, R. E. Crandall, y una especialista en ciencias de computación, Anna M. Penk, trabajando con un programa puesto a punto en el Centro de Cálculo del Reed College (Portland, Oregón), han conseguido probar que este número es efectivamente primo.

En el número de octubre mencioné un cuento de ciencia ficción, del que no recordaba ni autor ni título, en el que un personaje caminaba por el interior de un anillo prismático retorcido, de forma que tuviera solamente una cara. Muchos lectores, demasiados para poder citarlos aquí, me recordaron que el cuento era "What Dead Men Tell", de Theodore Sturgeon, y que apareció en el número de noviembre de 1949 de *Astounding Science Fiction*. Todo el número era una fantástica broma. En una carta publicada el año anterior, un lector había comentado críticamente el número a que nos referimos. El editor de la revista, John Campbell, Jr., consiguió convencer a la mayoría de los escritores mencionados en la carta para que escribieran cuentos con los títulos en ella mencionados, y así, cuando el número salió a la calle, la carta resultó ser modelo de precognición. El cuento de Sturgeon ha sido reimpreso en la colección *Imagination Unlimited*, editado por Everett Bleiler y T. E. Dikty (Berkley Publishing Company, 1959).

Taller y laboratorio

El efecto Moiré, el caleidoscopio y otros juguetes ópticos de la época victoriana

Jearl Walker

Este mes vamos a hablar acerca de unos clásicos juguetes ópticos que aparecieron en el siglo pasado. Aunque algunos de ellos todavía se venden, la mayoría han caído en desuso, víctimas de los nuevos entretenimientos, pero resultan tan divertidos e instructivos que merecían haber tenido mejor suerte.

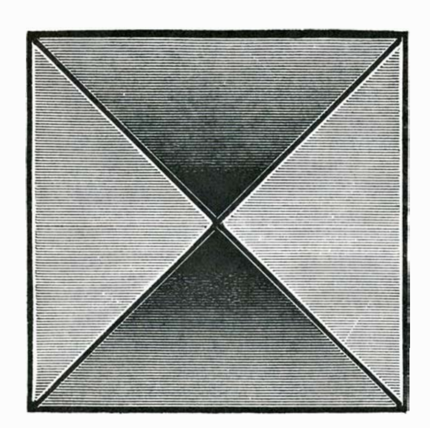
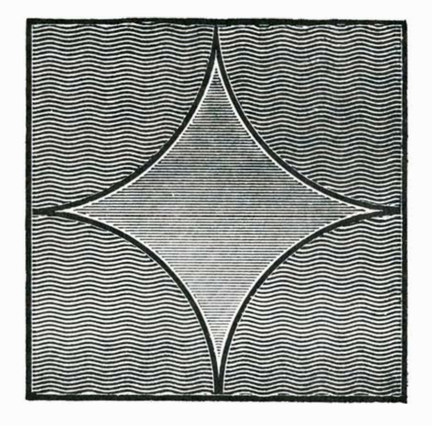
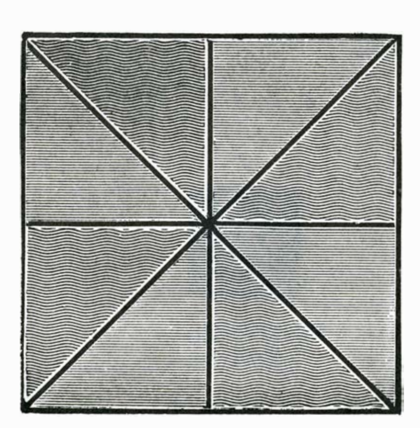
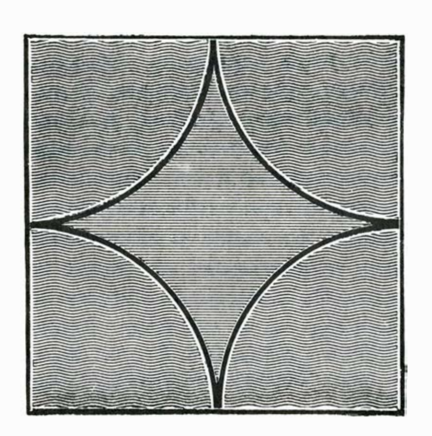
Uno de los juguetes ópticos más simples es el que ahora se llama "figuras moiré" (moiré patterns). Está basado en el efecto tremulante que se observa al mirar dibujos de figuras regulares, colocados unos sobre otros de manera que, en algunos sitios, las figuras coinciden (están en fase) y en otros no (están desfasadas). Se puede llevar a cabo una sencilla demostración de este efecto con un peine. Para ello se debe mantener éste paralelo a un espejo, de suerte que el observador pueda mirar a través de las púas y ver la imagen del peine en dicho espejo. Si el peine se coloca a una distancia adecuada con respecto al espejo, algunas púas de la imagen coincidirán con las reales en el campo de visión del observador. Por el contrario, otras púas estarán desfasadas con respecto a las reales. Lo que se ve es una configuración repetida a lo largo del peine que será obscura en las zonas en las que las púas reales no coincidan (estén desfasadas) con las de la imagen. Si se mueve el peine paralelamente al espejo, se podrá observar cómo las zonas obscuras (suponiendo que el peine sea obscuro) corren a lo largo del peine con un movimieto ondulatorio.

Este tipo de movimiento se empleó en dibujos publicados en el siglo XIX. Sobre un dibujo en blanco y negro se juntaron muchas líneas paralelas muy próximas unas de otras. Se pedía al lector que colocara sobre el dibujo un filtro, al que también se le habían pintado una serie uniforme de líneas paralelas. No estoy seguro de qué material estaba hecho el filtro, pero probablemente sería de cristal o algún tipo de papel transparente. Cuando el observador movía lentamente el filtro sobre el dibujo, la configuración rayada que originaba el conjunto de las líneas del filtro y la ilustración del dibujo, también rayada, daban la impresión de que los objetos del dibujo se estaban moviendo.

Dover Publications ha reimpreso (bajo el título de *The Magic Moving Picture Book*) un libro que se publicó por primera vez en Londres en 1898 con el título de *The Motograph Moving Picture Book*. El filtro transparente es de acetato y está rayado con una serie de líneas paralelas muy unidas. Los dibujos comprenden una serie de diseños geométricos, una representación de un volcán en erupción y una escena de un bombero apagando un incendio. Si movemos el filtro sobre el dibujo del bombero, por ejemplo, el movimiento de la configuración rayada hace que el fuego pierda fulgor, que aumente el humo y que el agua corra sobre el pavimento de la calle. En otras ilustraciones el efecto del rayado parece hacer que giren unas ruedas y el agua ondule en torno a un barco.

Se puede fabricar, con los propios medios, un duplicado de este efecto. Para ello se debe dibujar sobre una hoja de papel una figura geométrica con muchas líneas paralelas, separadas uniformemente. A continuación, se raya de forma similar, con líneas análogamente separadas, una lámina de acetato transparente. El plástico de acetato se puede adquirir en cualquier tienda de artículos de pintura artística, en donde también puede comprarse el tipo de tinta adecuado para dibujar sobre el plástico.

Una vez que la tinta se haya secado, se coloca el filtro sobre el dibujo y se mueve lentamente sobre él, de manera que las rayas del filtro y las del dibujo permanezcan paralelas. La configuración del dibujo se moverá de forma ondulante. Se pueden conseguir diversas variaciones en el juego. Para ello hay que trazar las líneas sobre el dibujo ligeramente más separadas o más juntas que las del filtro, o bien hacerlas onduladas en lugar de rectas. También se puede ensayar a orientar las líneas del dibujo con diversos ángulos con respecto a las líneas del plástico y realizar figuras compuestas de áreas con distintas orientaciones de las líneas. Por último, también se pueden dibujar escenas en las que sólo se muevan las zonas que interese.

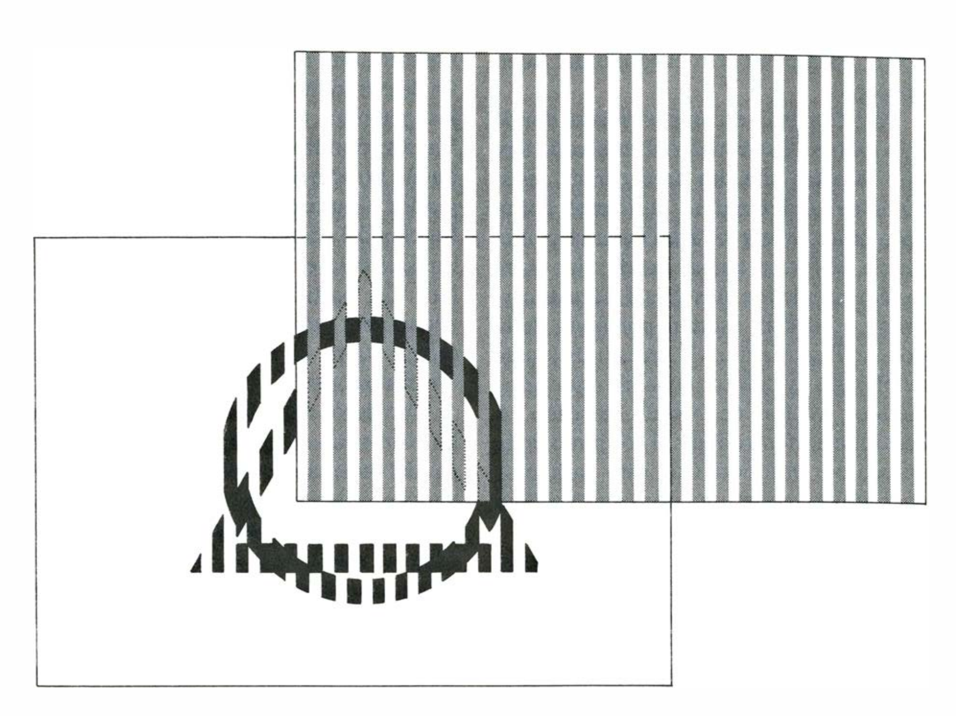
Diseños que pueden dar lugar al efecto moiré

Recientemente se ha utilizado un juguete similar por una cadena de supermercados de alimentación para promocionar una bebida. El juguete, ahora bajo patente a nombre de la empresa "Magic, Magic, Magic, Unlimited", consistía en dos tazas de plástico colocadas una dentro de la otra y que podían girar independientemente. La taza exterior, de plástico transparente, tenía una zona rayada con líneas negras verticales. La taza del interior era opaca y llevaba dos dibujos superpuestos alternándose uno con otro en espacios verticales del mismo ancho que las líneas que estaban pintadas en la taza exterior.

Con la taza interior en una determinada posición, las rayas negras verticales de la taza exterior tapaban las zonas verticales correspondientes a uno de los dibujos y dejaban ver, por el contrario, todas las partes correspondientes al otro dibujo. Desde una distancia de 30 centímetros más o menos, el observador ve el dibujo casi completo; las zonas negras son aparentemente ignoradas debido a que mentalmente se completa la figura rellenando los espacios ocultos por las líneas. Rotando la taza del interior se sustituyen las zonas visibles de uno de los dibujos por los del otro. Si se continúa rotando lentamente se van alternando una con otra las zonas visibles de los dos dibujos. En mi taza, un jugador de béisbol se encorvaba para detener una pelota o bien estaba de pie. Al girar la taza interior parecía alternarse entre las dos posturas.

La época victoriana se divertía con una gran variedad de juegos de dibujos intercambiables. El más simple de todos no consistía más que en un cuadro que tenía varias aberturas a través de las cuales podían deslizarse trozos del otro dibujo que iba a servir de alternativa. El movimiento cobraba mayor viveza cuando los dibujitos se hacían sobre dos plaquitas superpuestas que luego se proyectaban en la "linterna mágica", es decir, en lo que hoy llamamos proyector de diapositivas. El dibujo básico estaba pintado en una plaquita de cristal; la otra placa podía deslizarse frente a la primera. La segunda plaquita llevaba los elementos del dibujo compuesto que debían moverse. Cuando la "linterna mágica" proyectaba las dos plaquitas superpuestas, el operador sólo tenía que deslizar la segunda plaquita sobre la primera para conseguir dar el efecto de movimiento en la proyección.

Uno mismo se puede fabricar este sistema con la mayoría de los proyectores de diapositivas actuales. En el libro *Making Victorian Kinetic Toys*, Philip y Caroline Freeman Sayer recomiendan usar las delgadas laminitas de cristal de las monturas de las diapositivas de 35 milímetros. Para ello se puede comprar una pintura translúcida que se adhiera bien al cristal y que no se degrade con el calor del proyector. Pintar la escena fija sobre una de las laminitas de cristal. Es posible que se necesite preparar la segunda laminita para que se deslice mejor: para ello hay que ponerle una especie de asa, pegándole, por ejemplo, un palillo de dientes. De esa manera, cuando las dos laminitas estén metidas en la corredera del proyector, se puede mover la segunda lámina tirando del asa.



Diseños mezclados

Otro de los juguetes ópticos victorianos, el caleidoscopio, sigue gozando de la aceptación popular. Fue inventado (o, por lo menos, patentado) en 1816 por David Brewster, muy conocido por sus trabajos en óptica clásica, en particular por el llamado ángulo de Brewster en la polarización de la luz por reflexión. Aunque hay muchos tipos de caleidoscopios, el diseño básico es un tubo que contiene dos espejos o dos piezas de metal pulido pegados uno al otro formando un ángulo de 45 grados y con los otros dos bordes pegados a lo largo de la pared longitudinal del tubo. En un extremo hay un compartimento con trocitos de cristal o de cualquier otro material coloreado y en el otro un pequeño agujero a través del cual puede mirarse al interior del tubo. Mirando por dicho agujerito ve sólo un sector circular de todo el conjunto de los trocitos de material coloreado, así como las reflexiones de dicha zona en los espejos. Cuando el tubo se gira, la configuración simétrica que habíamos visto se va cambiando conforme los trocitos de material van cayendo en otras posiciones.

Se pueden encontrar instrucciones para montar un caleidoscopio casero en el excelente libro de Carson I. A. Ritchie, *Making Scientific Toys*. Lo primero que hay que hacer es recortar el fondo de una botella de plástico suficientemente larga y de las que tienen un gollete estrecho. La botella servirá como tubo y el gollete como apertura a través de la cual miraremos. Dentro del tubo se colocan dos espejos rectangulares con sus bordes formando un ángulo entre 40 y 70 grados. La botella deberá ser como un centímetro más larga que los espejos una vez que éstos se hallen colocados en su interior.

Colocar el fondo de la botella, una vez recortada, sobre una lámina de plexiglás de 3 milímetros de espesor y dibujar en ella la circunferencia de la botella. Hecho esto, se recorta el círculo de plexiglás trazado y se liman los bordes hasta que la pieza obtenida se ajuste perfectamente dentro de la botella.

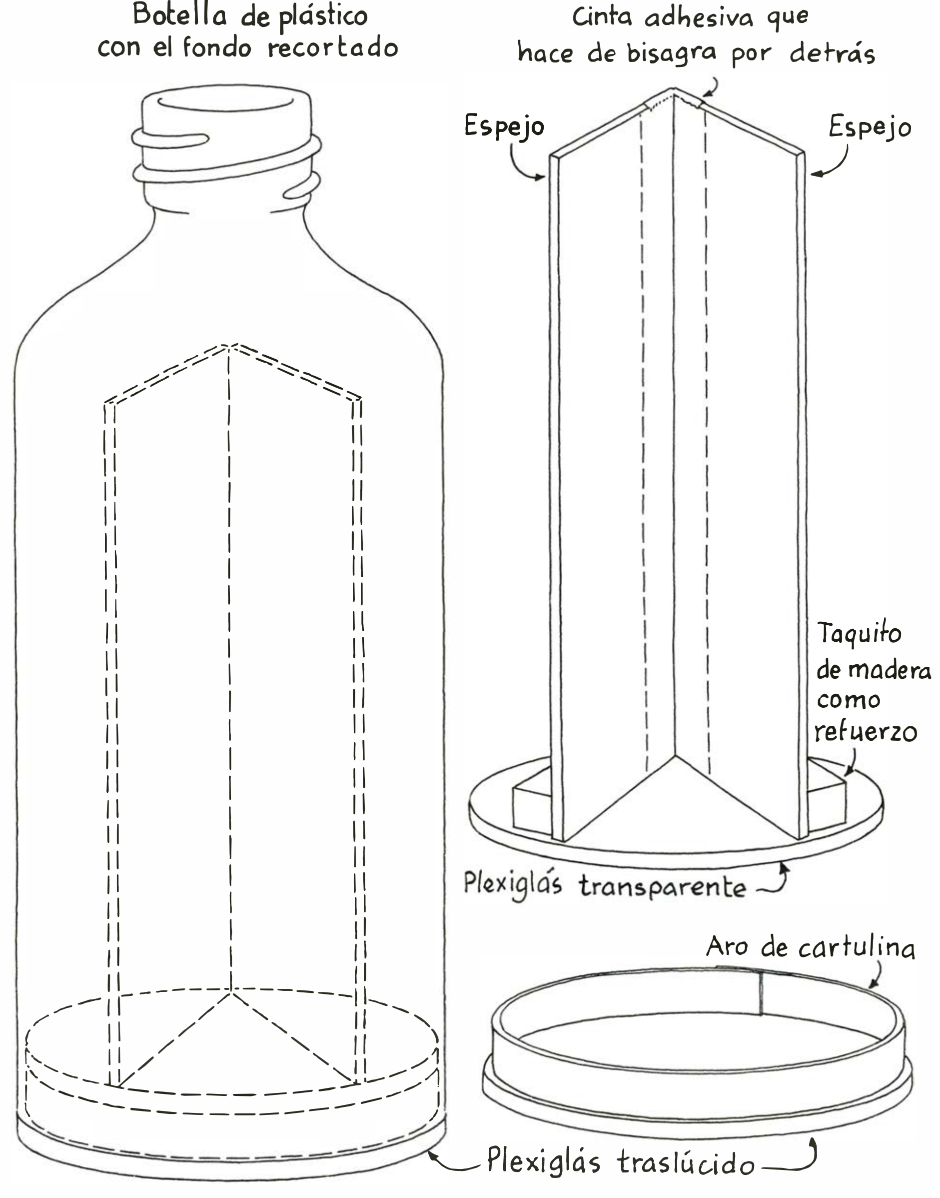
La unión en bisagra de los dos espejos se debe reforzar con cinta adhesiva. Luego se pondrá la pieza de plexiglás recortada sobre una mesa, se situará sobre ella perpendicularmente el conjunto de los dos espejos y se fijarán al plexiglás con pegamento, pero colocando además un trocito de madera, con idea de mantener fijo el ángulo entre los espejos. Se obtendrán mejores resultados si se colocan los espejos de manera que los dos bordes de cada uno de ellos lleguen a tocar la circunferencia del círculo de plástico. Cuando el pegamento empleado se haya secado, introdúzcase el conjunto en la botella tan profundo como sea posible, pero de manera que lo último a introducir sea el extremo con el redondel de plástico. Ahora haga otro círculo de plexiglás indéntico al anterior pero, para que no sea transparente y sí translúcido, ráspelo con virutas de

acero. Esta pieza será el fondo del caleidoscopio.

Para construir el compartimento donde irán los trocitos de material coloreado hay que recortar una tira de cartulina resistente de una longitud suficiente como para amoldarse a lo largo de la circunferencia de forma que se pueda fijar entre los dos círculos de plexiglás. La tira de cartulina no deberá asomar por fuera de la botella. Pegar la tira a la pieza translúcida de plexiglás y colocar en su interior unos cuantos trocitos de cristal coloreado. A continuación introducir el conjunto en la botella por el fondo. Con esto ya hemos conseguido construir un caleidoscopio con un muestrario intercambiable.

Para investigar cómo influye el ángulo de los espejos quizá sea preferible trabajar con un montaje más simple. A tal fin se puede hacer un dibujo geométrico en un papel y situar sobre el mismo perpendicularmente dos espejos colocados formando ángulo y mirar el dibujo anterior por la abertura que dejan los dos espejos. Para un determinado ángulo entre los dos espejos, ¿cuántos sectores circulares se ven? Vaya variando el ángulo desde el más pequeño posible hasta prácticamente los 90 grados y anote el número de imágenes en función del ángulo. ¿Tienen todos los sectores circulares la misma anchura angular? En particular, ¿el sector circular situado por detrás de la bisagra de los dos espejos tiene el mismo tamaño que las otras secciones?

Uno de los mayores caleidoscopios que he visto tenía en el fondo una lente en lugar del receptáculo (creo que el aparato se llama propiamente teleidoscopio, pero no estoy seguro). La distancia focal de la lente debe ser menor que la longitud del tubo. Cuando se mire a través del aparato un objeto convenientemente separado, la lente reduce el tamaño de su imagen, de manera que se ven a la vez muchos objetos iguales. Para mí el efecto es maravilloso debido a que hay una mayor luz y un mayor colorido que en las típicas imágenes de un caleidoscopio convencional, donde los trocitos de cristal coloreados, así como el fondo translúcido, reducen la intensidad de la luz. Si se mira con el aparato un paisaje en movimiento, por ejemplo yendo en un coche, las constantes variaciones de las imágenes son casi sicodélicas.



Diseño de Carson I. A. Ritchie de un caleidoscopio casero

Uno se puede fabricar un caleidoscopio de este tipo con una lente barata, un tubo de cartulina resistente y los dos típicos espejos. Elegir la lente y el tamaño del tubo de manera que el foco de la lente quede hacia la mitad del tubo. La abertura del extremo para mirar debe ser pequeña, pero lo suficientemente grande como para poder contemplar toda la imagen, incluyendo las reflexiones en los espejos.

Otra diversión, enormemente popular en el siglo pasado, consistía en mirar fotografías en estereoscope. Ritchie en su libro da normas acerca de cómo hacer una réplica de los viejos estereoscopios. En los anticuarios se pueden encontrar viejos aparatos de éstos y postales estereoscópicas, incluso se pueden comprar repeticiones de las mismas en el Stereo Classics Studio, Inc., 145 Algonquin Parkway, Whippany N.J. 07981. El tipo más común de estereoscopio tiene dos lentes, una para cada ojo, a través de las cuales el observador ve dos fotografías, una por cada ojo. Las fotos no son idénticas, sino que dan una perspectiva ligeramente diferente de la misma escena de manera que, al unirse las perspectivas de cada ojo, pueda verse la misma composición. Cuando las fotos se funden en el estereoscopio, se ve el original en tres dimensiones.

En su libro *Ready to Make Photo & Scene Machines*, Budd Wentz describe unos procedimientos muy sencillos para fabricar diferentes tipos de estereoscopios. El observador mira a través de una pieza de cartulina parecida a una máscara Halloween. Uno de los fines de la máscara es eliminar las distracciones de manera que uno se pueda concentrar en todos los detalles que le ayudarán a situar cada objeto de la fotografía en la profundidad que le corresponda.

En el centro de la máscara hay un trozo de cartón duro que tiene pegado un espejo plano. Se trata de ver una imagen directamente con un ojo mientras se ve una reflexión de la otra en el espejo con el otro ojo. Este sistema hace que los ojos converjan, lo cual, a su vez, ayuda al observador a dar profundidad a las figuras. Como ocurría antes, las dos figuras muestran perspectivas ligeramente diferentes; pero, además, una de ellas debe cambiarse de posición (volviéndola al revés) de manera que tenga la misma orientación que la otra después de la reflexión. Si se hace todo bien, el observador funde en una ambas

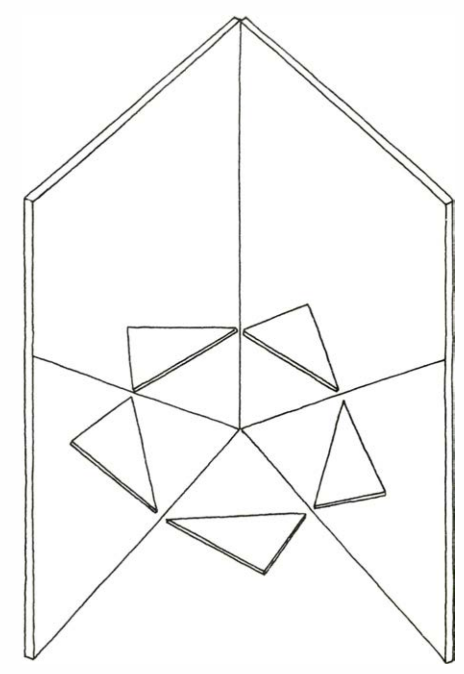
imágenes para obtener una sola tridimensional. Wentz proporciona reproducciones de algunas de las tarjetas estéreo que fueron populares con este tipo de aparato a finales del siglo XIX y principios del XX.

El estereoscopio corriente que se puede adquirir en las tiendas de antigüedades es diferente del de Wentz: tiene dos lentes en forma de prisma para facilitar la mezcla de las distintas imágenes que ve el observador. Ambas imágenes poseen la misma orientación, ya que no hay reflexión. Este tipo de estereoscopio parece que fue creado en 1859 por Oliver Wendell Holmes, basándose en parte en trabajos anteriores de Brewster. Las fotografías se hacían normalmente con una cámara provista de dos lentes, las cuales estaban separadas entre sí por una distancia aproximada a la separación media existente entre los ojos de una persona. Las distancias focales de las lentes en la cámara y en el estereoscopio eran las mismas, de manera que el observador veía la misma perspectiva que la cámara. También se podían hacer las fotos con una cámara ordinaria tomando primero una fotografía y moviendo luego la cámara siete u ocho centímetros hacia un lado para hacer la otra.

Thomas B. Greenslade, Jr., y Merritt W. Green III han publicado el resultado de sus experimentos con este tipo de estereoscopio y sus intentos de hacer fotografías para el mismo. Es posible que el lector quiera seguir sus pasos bien con un estereoscopio antiguo o con uno casero. Si el lector se construye uno en casa, Ritchie sugiere que se encarguen las lentes a un óptico. Se podría probar a utilizar en su lugar prismas de vidrio en forma de cuña.

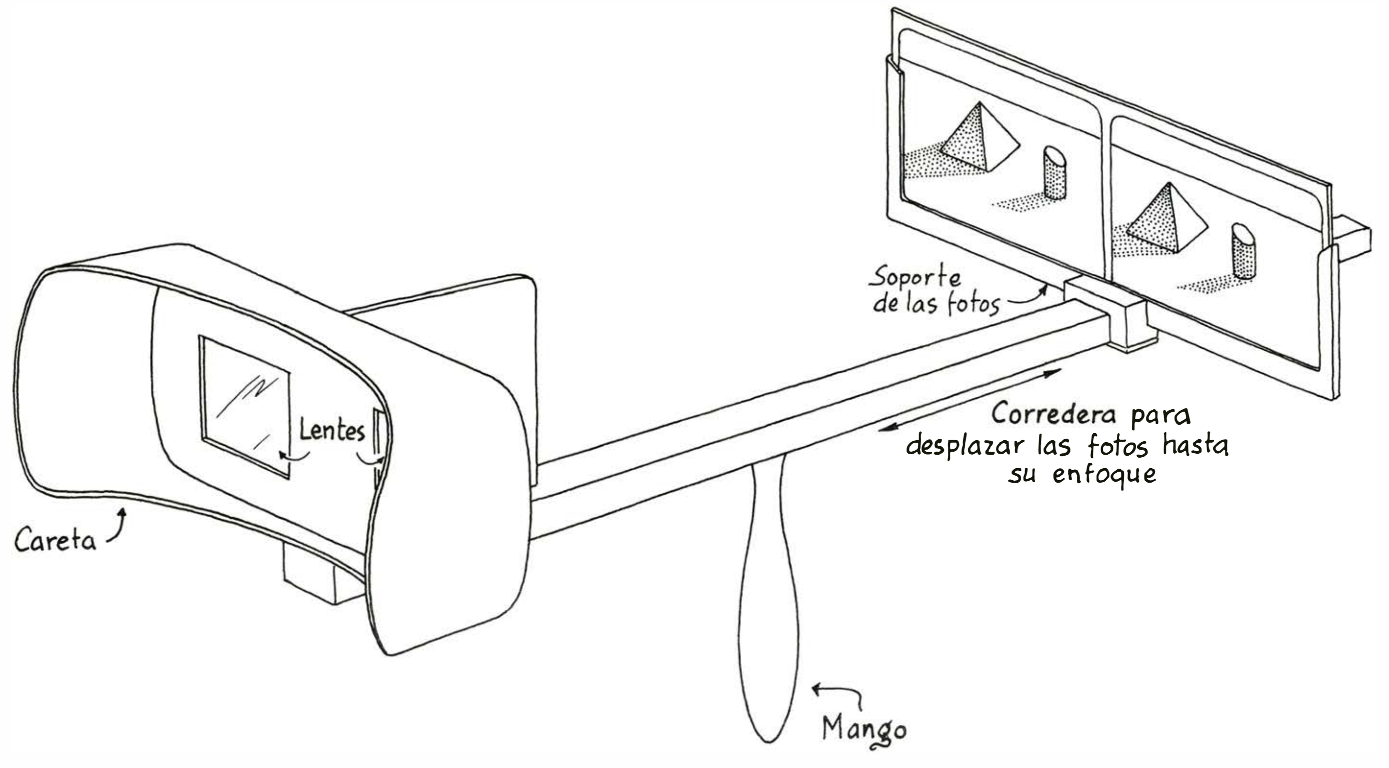
Greenslade y Green fueron capaces de hacer sus fotografías con una cámara Polaroid (Modelo 180) que tenía una lente con distancia focal parecida a la del antiguo estereoscopio que habían comprado. En un experimento querían confirmar cómo el desplazamiento lateral de la posición de la cámara entre sucesivas fotografías afectaba a la calidad de la ilusión tridimensional. Idealmente el desplazamiento debería ajustarse a la distancia entre los ojos del observador, pero la profundidad aparecía en las fotografías incluso cuando el desplazamiento de la cámara era mayor o menor que la distancia ideal. Con un desplazamiento grande, sin embargo, los objetos en la fotografía aparecían irreales en proporción de delante atrás debido a que la mayor separación les daba demasiada profundidad.

Greenslade y Green encontraron también que pueden hacerse fotografías estéreo rotando la cámara un pequeño ángulo (7,7 grados) entre fotografías sucesivas. La rotación parecía alterar la perspectiva suficientemente como para permitir que el observador percibiera la profundidad en las imágenes al mezclar éstas. Es posible que el lector desee experimentar con un rango de ángulos para ver si el efecto tridimensional persiste y si ángulos amplios dan lugar a distorsiones en la imagen percibida.

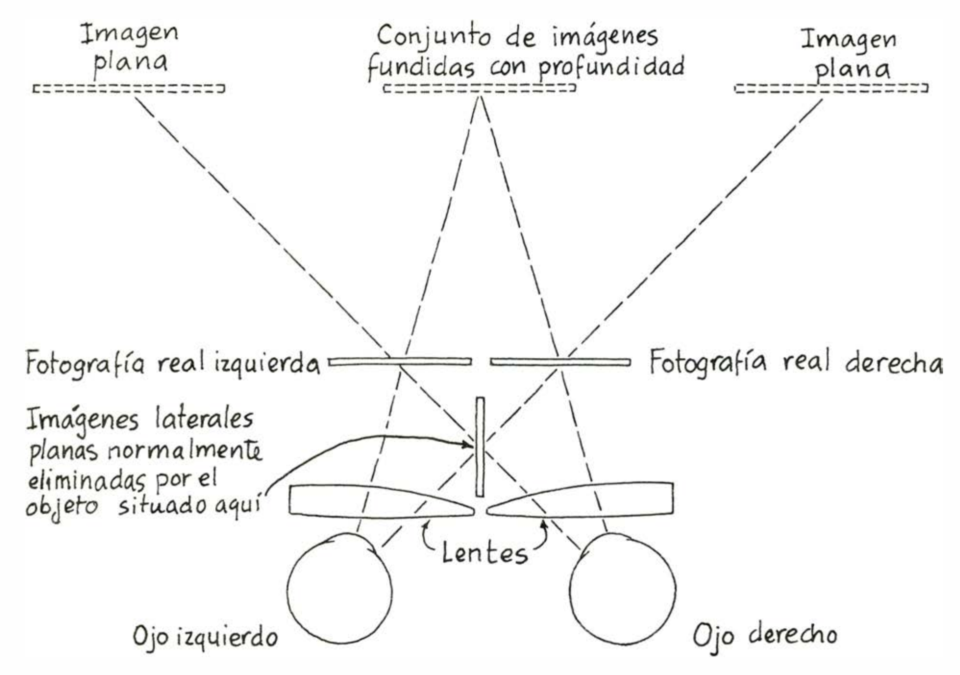


Un modelo de reflexión múltiple

A mí me resultó bastante fácil, usando una Polaroid modelo 104, hacer fotografías estéreo de escenas en la calle sujetando la cámara con las manos. Si se usa un trípode para desplazar o girar la cámara se mejorará, sin duda, la calidad de las fotografías, pero la ilusión tridimensional aparecía con facilidad, incluso a pesar de mi escaso equipo. Enfoqué la parte izquierda del campo de visión de la cámara a un objeto e hice una fotografía. Después giré la cámara unos 5 o 10 grados hacia la izquierda, la enfoqué de forma que el objeto de referencia



Componentes de un estereoscopio



Optica de un estereoscopio

estuviera a la misma altura en el campo de visión e hice la segunda fotografía.

Luego recorté los bordes blancos de las fotografías y las puse en el estereoscopio. Al mirar las fotografías a través del instrumento, desplacé una de ellas a través de mi campo de visión hasta que se acercó a la posición correcta, después las fotografías se juntaron de repente para componer una sola imagen con profundidad. Entonces las pegué y las monté sobre un trozo de cartón para mantenerlas con la separación correcta.

Cuando las fotografías ya estaban colocadas correctamente, mi ojo izquierdo veía más de la parte izquierda de la escena que mi ojo derecho. Esta diferencia es automática si se hace girar o se desplaza lateralmente la cámara. Incluso fotografiando exactamente la misma escena sin mover la cámara, entre dos exposiciones sucesivas, conseguí obtener una buena ilusión tridimensional manipulando la posición de las fotografías en el estereoscopio. Sin embargo, esto no siempre es posible.

A mí me parece francamente divertido el fabricarse uno mismo las propias tarjetas estéreo. Se puede usar incluso una vieja instantánea Polaroid si se tiene duplicada, bien por Polaroid o haciéndose uno mismo una fotografía. Para conseguir nuevas parejas de fotografías el mejor tipo de escenas a captar es el que tenga objetos tanto en primer plano como en segundo.

Normalmente se asigna la profundidad a una escena con la ayuda de algunas pistas, entre ellas el hecho de que se conoce el tamaño relativo de los objetos que se ven y, por tanto, se puede calcular su distancia en parte con los ángulos relativos que subtienden en el campo de visión. Asimismo ayudan las sombras y la geometría convergente de objetos tales como paredes y carreteras, cuando se las ve alejándose del observador.

También se siente inconscientemente la acomodación y convergencia de los propios ojos cuando se mira un objeto en una escena. La acomodación refleja la cantidad de tensión muscular que debe ejercerse sobre la lente del cristalino del ojo para alterar su forma de manera que la imagen del objeto esté enfocada en la retina. Esta pista para la profundidad no depende de un juego entre los dos ojos, sino que es monocular.

La convergencia es una pista binocular relacionada con el ángulo en los ejes ópticos de los ojos. El eje es una línea imaginaria que pasaría por el centro de la lente y penetraría en la fóvea, que es la pequeña depresión de la retina donde se concentran los fotorreceptores. La visión normal supone el movimiento de los ojos hacia la izquierda o derecha hasta que la imagen del objeto queda dentro del campo de la fóvea. Para un objeto distante, los ejes ópticos de los dos ojos son casi paralelos. Para un objeto próximo forman un ángulo acusado. El control muscular necesario para hacer que los ejes ópticos converjan en un objeto ayuda a calcular la distancia a que se encuentra.

Supongamos que el observador ve una colección de objetos que están a distintas distancias. Tendrá una idea de sus diferentes profundidades aun cuando no se los mire individualmente porque las posiciones de sus imágenes en la retina alrededor de lá fóvea serán diferentes para los dos ojos simplemente debido a la separación de éstos. La diferencia de posición con respecto a la fóvea será mayor con un objeto próximo que con uno distante. A través de este efecto, el cerebro puede asignar las profundidades distanciadas adecuadas a los objetos.

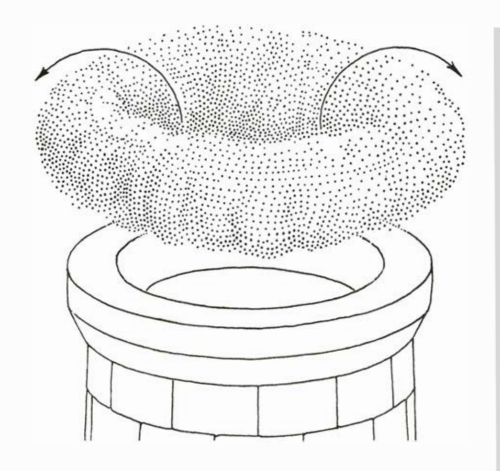
Para volver a crear la profundidad en una fotografía bidimensional, el problema a solucionar será conseguir que el cerebro ignore las pistas que sugieren la bidimensionalidad al ver la fotografía. Cuando se mira ésta, tanto el enfoque de cada ojo como la convergencia entre ellos nos informan que estamos mirando una fotografía y no una escena real. Si se colocan dos fotografías adyacentes, el efecto es el mismo, el cerebro sabe que se están mirando dos fotografías adyacentes.

El estereoscopio nos permite enfocar las fotografías mientras se está obligando a los ojos a converger; resulta así que las imágenes ocupan en cada retina las mismas posiciones que ocuparían si se estuviera mirando a la escena real. Esta sensación es tan fuerte, al menos para la gente que puede ver la ilusión tridimensional, que uno se convence de que hay profundidad en las fotografías mezcladas. De esta forma el cerebro está tan dispuesto a asignarles profundidad que las fotografías en el estereoscopio no tienen que estar exactamente en su posición correcta. A medida que se acercan a la posición adecuada el cerebro aúna de inmediato sus imágenes y empieza a asignarles profundidad. Esto puede hacerlo en cierta medida incluso con fotografías que son duplicados exactos.

Aun cuando un estereoscopio es un medio cómodo de obtener esta ilusión, no es estrictamente necesario. Se puede ver profundidad en una simple tarjeta estéreo, por medio de un simple truco. Después de fijar los ojos en un objeto que esté situado a unos 150 centímetros deslice rápidamente una tarjeta estéreo en el campo de visión. Mientras lo va enfocando, trate de mantener la misma convergencia de los ojos que tenía para el objeto alejado. Con alguna práctica este truco le permitirá mezclar las dos imágenes de la tarjeta estéreo en una sola con profundidad, lo mismo que se hace con un estereoscopio. Yo mismo puedo crear esta ilusión tridimensional a voluntad y sin el truco descrito, simplemente mirando con fijeza una tarjeta estéreo mientras ajusto la convergencia de mis ojos hasta que las dos imágenes, una de cada ojo, se funden.

Tanto si se hace por medio del truco como si se hace con la técnica de la fuerza bruta, lo que se ve es notablemente distinto de la ilusión del estereos-

copio. Sin el instrumento se ve, en cada lado de la imagen con profundidad, otra imagen adicional bidimensional de la escena en la tarjeta estéreo. Greenslade y Green ofrecen una demostración para explicar las dos imágenes planas. Se deben fijar los ojos en un objeto colocado a una distancia de unos 150 centímetros, sujetando simultáneamente dos lápices en posicion vertical a unos 25 centímetros delante de los ojos separados entre sí unos 8 centímetros. Se observará que no se ven las imágenes de dos lapiceros, sino de cuatro, dos en cada ojo. A medida que uno se concentra en el objeto distante, se deben ir acercando los lápices entre sí hasta que se mezclan las dos imágenes más próximas. Esta mezcla de imágenes es la que forma la tridimensional cuando se mira una tarjeta estéreo a través del estereoscopio. Con los lapiceros las dos imágenes exteriores están todavía presentes: la de la izquierda corresponde al lápiz de la izquierda visto por el ojo derecho y la imagen de la derecha es la del lápiz derecho visto por el ojo izquierdo. En el estereoscopio esas dos imágenes exteriores que distraen se eliminan por medio de la fina pared del aparato, colocada entre los ojos y perpendicularmente a la cara.



Un remolino de humo anular

En julio escribí sobre la forma de los penachos de humo de las chimeneas bajo varios perfiles de temperaturas verticales. G. Frederick Stork de Chevy Chase, Md., me ha enviado una fotografía de un hecho raro: un anillo de humo. Aun cuando él no vio la formación del mismo, ambos creemos que éste debió ser expulsado por la chimenea que aparece en primer término. Cuando se envía una bocanada de gas caliente a través de una abertura circular, la viscosidad entre la pared y los gases adyacentes a la misma hacen que esos gases se muevan más lentamente que los del centro, consiguiendo que los últimos se ondulen sobre los primeros una vez que han emergido. Esta ondulación crea la estructura anular. El cañón de la chimenea debe haber condensado de alguna manera el agua de la bocanada y así el anillo se hace visible.



Fotografía de un anillo de humo de una chimenea, realizada por G. F. Stork

Libros

Tres textos universitarios sobre mecanismos y máquinas, cristalografía y cirugía pediátrica

Lorenzo Alvarez, Juan A. Campá y José Argemí

SÍNTESIS DE MECANISMOS, por Justo Nieto Nieto, Editorial AC. Madrid, 1978, 190 páginas. Presentación por el catedrático de cinemática y dinámica de máquinas de la Escuela Técnica Superior de Ingenieros Industriales de Madrid, Emilio Bautista Paz. Desde los tiempos más remotos, el hombre ha utilizado la máquina con el doble fin de aliviar su trabajo y hacerlo más noble y humano. Por ello, la humanidad siempre ha intentado, intenta e intentará, perfeccionar las máquinas ya existentes, o diseñar y construir otras de distinta concepción que permitan disminuir o eliminar el trabajo menos noble, el muscular, por el más noble, el intelectual; así como lograr máquinas más perfectas, de mayor y mejor calidad de producción. El logro de tales máquinas ha permitido, a los países más adelantados, llegar a unos niveles de calidad de vida que de otra forma hubiese sido imposible conseguir.

A mi modo de ver, la estructura dada por el autor a la obra, en seis capítulos, es acertada. En el primero, en forma de introducción, se hace un pequeño resumen de la historia de los mecanismos y máquinas, sobre la síntesis de mecanismos y su desarrollo actual. Los cuatro siguientes están dedicados al estudio y exposición de cada uno de los criterios existentes actualmente en la síntesis de mecanismos; y el último a la optimización en síntesis.

La obra constituye una admirable y excelente labor de recopilación, clasificación y ordenación de la principal y múltiple información dispersa que sobre análisis, síntesis, diseño y optimización de mecanismos existe actualmente, y que, en una extensa bibliografía, se citan en las últimas páginas del libro. La mayoría de esta bibliografía procede de los EE. UU. y de Rusia, aunque no faltan la alemana, francesa, inglesa, mexicana, rumana y española del propio autor de esta obra.

En el desarrollo de las máquinas, se pueden distinguir actualmente las siguientes etapas: *a*) máquinas de la primera generación, caracterizadas por el empleo de mecanismos de concepción simple y materiales poco desarrollados; ello daba lugar a máquinas de baja capacidad de producción; estas máquinas se emplearon hasta el inicio de la revolución industrial. *b*) Máquinas de la segunda generación, caracterizadas por la utilización de mecanismos de concepción más compleja y utilizándose en ella materiales de mejor calidad y más variados, lo que permitió velocidades más elevadas, mayor capacidad de producción y calidad más óptima; estas máquinas hicieron posible la revolución industrial. *c*) Máquinas de la tercera generación, caracterizadas por mecanismos muy perfeccionados, perfectamente equilibrados, que han permitido el logro de grandes velocidades. El desarrollo paralelo de otras tecnologías ha permitido integrar en estas máquinas aparatos de control, regulación y automación de concepción no mecánica. Son éstas las máquinas de la actualidad.

Los conceptos mecanismo y máquina son completamente diferentes. La integración más o menos afortunada de uno o más mecanismos constituye lo que universalmente se ha aceptado en denominar "esqueleto geométrico cinemático" de una máquina. En el diseño de una máquina intervienen, además de una correcta concepción de este esqueleto geométrico-cinemático que asegura su funcionalidad, otros factores de tipo socio-político-cultural que inciden y configuran el proceso de distribución-consumo, que de ignorarse pueden dar lugar a un diseño erróneo y poco competitivo, por exceso de precio, tamaño o peso, por falta de estética, por peligroso funcionamiento, mantenimiento o reparación, por elevado nivel de vibraciones o ruidos, etcétera. El proyectista de una máquina debe adoptar una postura crítica que tenga presente múltiples factores, unos de tipo mecánicos y otros de tipo amecánicos. Entre los primeros, es, sin dudar el más importante, la función exigida a la máquina, y ésta se hallará mejor realizada cuanto más elevados sean los conocimientos que tenga el proyectista de los mecanismos existentes, de sus múltiples combinaciones (síntesis estructural); de la coordinación de las posiciones de las barras de entrada y salida de un mecanismo (síntesis de generación de funciones); de la correspondencia de las trayectorias descritas por puntos pertenecientes a las barras de un mecanismo, durante el movimiento de éste con otras trayectorias especificadas (síntesis de generación de trayectorias); de la posibilidad de situar una determinada barra del mecanismo en diversas posiciones especificadas (síntesis de guiado del cuerpo rígido); de la elección de entre las múltiples posibles, la mejor y más simple (optimización de síntesis).

En su primera parte "Sobre la historia de los mecanismos y máquinas", el autor, con el fin de situar al estudioso, expone someramente en once páginas la evolución de la teoría de mecanismos y máquinas, desde las más antiguas referencias escritas, encontradas en tres poemas de la literatura india (1700 a. C.), en las que aparecen menciones sobre carros y ruedas, hasta nuestros días, con el fecundo desarrollo de las modernas escuelas norteamericana y rusa, así como otras muy activas como la australiana, austríaca, búlgara, canadiense, checoslovaca, francesa, holandesa, india, inglesa, italiana, japonesa, polaca, rumana y yugoslava.

Todo ello dividido en tres grandes períodos, desde la antigüedad hasta finales del siglo XIX, en que se acumulan muchos datos como consecuencia de la creatividad semiempírica; en él nace la primera escuela sobre la teoría de mecanismos y máquinas, iniciada en Francia por Monge (1746-1818), célebre matemático francés, notable por sus estudios geométricos y uno de los fundadores de la Escuela Politécnica de París. En este primer período se citan las aportaciones tanto técnicas como bibliográficas, sobre mecanismos y máquinas, de Aristóteles, Arquímedes, Vitruvius, Herón de Alejandría, Jordán, Leonardo da Vinci, Euler, Watt, Betancourt, Willis y Ampère entre otros. En él se establece la primera clasificación de los mecanismos en orden a la transformación del movimiento, hecha en 1810 por Hachett en su obra *Tratado elemental de máquinas,* así como la clasificación de Borgnis en 1818 que divide la máquina en seis partes (receptores, comunicadores, modificadores, soportes, reguladores y operadores), la establecida por Willis en 1841 en su libro *Principles of Mechanisms,* que clasifica los mecanismos en tres clases y en tres géneros. Los conceptos de par

cinemático, y su división en superiores e inferiores, el de barra y cadena cinemática, los de inversiones y equivalencia cinemática, de tan fecunda aplicación, fueron establecidos por Reauleaux en 1875 en su texto pionero *Cinemática*. Este período termina con la iniciación de dos grandes escuelas: la alemana de Reauleaux y la rusa de Chebyshev.

En el segundo período se produce un desarrollo espectacular de la teoría de mecanismos y máquinas en las escuelas alemana y rusa. Burmester desarrolla la escuela geométrica, aplica por primera vez los polígonos de velocidades y aceleraciones, extiende las ideas geométricas a la síntesis dimensional de mecanismos planos fundamentalmente. Esta síntesis concierne al posicionado de un punto en una determinada trayectoria, al guiado de un plano, y en ella se utilizan conceptos de la geometría plana tales como la ecuación de Euler-Savary, circunferencias de Bresse, puntos de Burmester, cúbica de curvatura estacionaria, punto de Boll, etcétera. En Rusia, Chebyshev y sus discípulos se especializan principalmente en las técnicas de ajuste y aproximación de curvas, desarrollando técnicas especiales y aplicando nuevas herramientas matemáticas a la síntesis de mecanismos. Bloch introduce los números complejos, Dimentberg utiliza los números duales en el análisis de mecanismos espaciales y Kislitsin introduce el cálculo tensorial. Aunque en este período van a la cabeza las dos escuelas ya citadas, la alemana y la rusa, no por ello debe olvidarse las aportaciones acaecidas en Australia por Rosenauer, Willis, Hirschhorn y Hunt; en Austria por Poeschl y Federhofer; en Francia por Nicaise, Koenings, Bricard y Manheim; en Holanda por Bottema; en Italia por Allievi; y en Rumania por Manolescu, Pelecudi y Mangeron.

El tercer período se inicia con la incorporación de los EE. UU. alrededor de 1940, en el estudio de la síntesis de mecanismos. Hartemberg y Denavit se lamentaban en dicho año del escaso interés que había en este país por la síntesis de mecanismos, y de Jonge afirmaba, en 1942, que los Estados Unidos no había participado en el desarrollo de esta rama del saber. La escuela norteamericana nace como consecuencia de las investigaciones llevadas a cabo casi simultáneamente por: Adams, Beskol, Crosley, Davies, Hall, Johnson, Lowen, Raichel, Shigley entre otros. En la dos últimas décadas, la teoría de mecanismos y máquinas ha experimentado un gran desarrollo como consecuencia de la creación del organismo International Federation for the Theory of Machines and Mechanisms; se han desarrollado congresos internacionales y nacionales, se ha facilitado el intercambio de puntos de vista entre investigadores de este campo del saber y han nacido nuevas revistas técnicas dedicadas exclusivamente a esta materia, así como se han escrito y publicado gran número de artículos en otras revistas.

Se conoce como "Síntesis estructural" o "Síntesis de mecanismos", la parte de la síntesis que trata de la selección de tipos de mecanismos y de las diversas maneras de producirse la formación de las cadenas cinemáticas y de los mecanismos. Esta síntesis consiste, por tanto, en concebir, a partir de una normativa propia y para exigencias primarias de diseño, los mecanismos que la cumplen.

El doctor Nieto Nieto desarrolla esta síntesis en el capítulo segundo de la obra; lo efectúa con una adecuada ordenación, de forma tal que pasa desde los conceptos más elementales a las leyes más complejas de formación de las cadenas cinemáticas planas. Inicia este estudio con la definición de barras en sus distintas modalidades de: rígidas, unirrígidas, elásticas y fluídicas; de par cinemático y cadena cinemática, mecanismos y sus grados de libertad y movilidad. En el estudio de los pares cinemáticos, éstos se encuentran clasificados en cinco clases, según el número de grados de libertad del par sea de uno a cinco, siendo lo más acertado de todo ello el cuadro de clasificación de los mismos que inserta, en el que se incluye un esquema del par. Seguidamente, estudia la obtención de mecanismos por la adición de diadas de rotación y grupos de Assur, de diadas prismáticas helicoidales, resortes y par cilíndrico, por adición de elementos unirrígidos y de pares de levas. La equivalencia cinemática la expone breve y claramente, así como la obtención de mecanismos por degeneración cinemática e inversión cinemática. Completa este capítulo con una demostración de las leyes de Grashof y de las leyes de formación de las cadenas cinemáticas planas, cuyo estudio topológico-estructural ha interesado a gran número de investigadores en la síntesis de mecanismos. Estas leyes permiten dar respuestas a preguntas tales como: *a)* ¿Cuántas cadenas cinemáticas de movilidad dada pueden formarse con un número especificado de barras? *b)* ¿Cuántas cadenas cinemáticas, isomorfas entre sí, existen para cada configuración de barras y pares?

En su forma general, es decir, para cadenas cinemáticas especiales de dimensiones cualesquiera el problema no ha encontrado solución, pero sí lo ha logrado para cadenas cinemáticas planas con pares de rotación, prismáticos, helicoidales y levas. Los métodos seguidos por los diversos investigadores son muy variados, desde los menos científicos, los intuitivos, a los más racionales con el empleo de matrices. El método más utilizado para resolver tales problemas es el de Grübler-Chebyshev, que se expone en la obra, así como otro muy reciente del autor de la misma, aplicable a cadenas planas con pares de rotación. El capítulo finaliza con una serie de problemas resueltos que ayudan a comprender todo lo expuesto.

"Sintaxis de Generación de Funciones" estudia la coordinación de las barras de entrada y salida de un mecanismo. En este tipo de síntesis se presenta uno de los siguientes problemas: *a)* Diseñar un mecanismo que, en el intervalo de su movimiento, sea capaz de situar las barras de entrada y de salida en dos grupos de posiciones especificadas. *b)* Diseñar un mecanismo tal que la función generada por él, o sea, la relación funcional de entrada y salida, coincida con una función, en un número de puntos especificados sobre ésta.

El autor estudia este tipo de síntesis en el tercer capítulo, y lo inicia con una introducción en la que expone las ideas fundamentales sobre la misma, para continuar con la ecuación de Freudenstein, aplicable al cuadrilátero articulado o a mecanismos degenerados del mismo, como, por ejemplo, el de biela-manivela. Esta ecuación permite resolver el problema de la síntesis de mecanismos con tres puntos de precisión. Si se desean mayor número de puntos de precisión, se puede introducir en la ecuación de Freudenstein nuevos parámetros de diseño que hagan el papel de incógnitas, y a continuación se especifican tantos puntos de precisión como ecuaciones sean necesarias para que el sistema sea determinado. Este sistema tiene la dificultad de no ser lineal, y exige programas y métodos de resolución que pueden ser costosos. Procediendo de esta forma, el autor analiza los casos de cuatro, cinco y seis puntos de precisión. Si se desea mayor aproximación en algún punto de precisión, se puede poner la condición adicional de que ambas funciones, la generada y la deseada, tengan en algunos de estos puntos de precisión, las primeras, las segundas, etcétera, derivadas iguales. Este tipo de síntesis se conoce por el nombre de "Síntesis con derivadas de precisión".

Evidentemente cada oposición de coincidencia en las derivadas equivale a una nueva condición o ecuación adicional y comoquiera que el número de

parámetros de diseño es constante, el número de puntos de precisión se ve disminuído en el número de derivadas que se desee hacer coincidir. Las ecuaciones adicionales se obtienen de las derivadas sucesivas de la ecuación de Freudenstein.

La síntesis de Bloch utiliza las derivadas temporales primera y segunda de las tres de barras móviles del mecanismo de cuatro barras o de sus degenerados, lo que permite determinar las dimensiones de las barras y la posición del mecanismo en la que las velocidades y aceleraciones angulares de las barras móviles tienen valores pre-establecidos.

Después del estudio analítico de generación de funciones, en la obra se exponen los métodos gráficos para resolver tales problemas. El autor estudia los siguientes cuatro métodos: el del polo, el del polo relativo, el de la inversión, y el de Hain o de reducción de puntos de precisión.

La "Síntesis de generación de trayectorias" es la parte de la síntesis de mecanismos que estudia la correspondencia de las trayectorias descritas por puntos pertenecientes a las barras de un mecanismo, durante el movimiento de éste, con otras trayectorias especificadas. Esta síntesis tiene por objeto obtener un mecanismo tal que a un punto de una determinada barra, generalmente un acoplador o barra flotante, se le impongan unas determinadas exigencias, tales como generar una trayectoria exacta dada, o aproximada, simpétrica respecto a un eje, con puntos dobles, con puntos de retroceso, con tramos casi circulares o rectilíneos, que pase durante su movimiento por un número fijado de puntos de precisión de una trayectoria dada, o que genere la misma curva que otro mecanismo dado.

Esta síntesis se expone con gran rigor en el capítulo cuarto, y ella se inicia con el estudio general de la trayectoria de un punto del acoplador del cuadrilátero articulado. Esta trayectoria depende de nueve parámetros, y el autor expone las principales propiedades de la misma, tales como la carencia de puntos angulares, si existen puntos dobles éstos se presentan donde la trayectoria corta a la circunferencia de los focos, la inexistencia de puntos de multiplicidad superior a dos, y la existencia de cúspides.

La generación de trayectorias puede realizarse con nueve puntos de precisión como máximo. A este respecto el doctor Nieto Nieto indica la dificultad de resolución de un sistema no lineal de nueve ecuaciones, el que prácticamente sólo puede lograrse con un ordenador, así como que algunas de las soluciones encontradas sean irrealizables por ser imaginarias. Hemos de hacer constar la elegancia al exponer la resolución del problema con tres puntos de precisión mediante el uso de números complejos, para el que indica un programa en Fortran V.

Una vez establecida la base científica de la síntesis de generación de trayectorias, el autor completa este cuarto capítulo con la exposición de una serie de métodos gráficos para la síntesis de generación de trayectorias, los que son de aplicación a casos especiales, eliminan las dificultades analíticas ya expuestas y dan una visión no abstracta del problema. Finalmente se estudia la teoría de la curvatura en el movimiento plano de un sólido rígido, con el fin principal de poder introducir el concepto de 'la cúbica de curvatura estacionaria y la posibilidad de síntesis que ofrecen ciertos puntos de la misma.

La "Síntesis de guiado de biela", o síntesis de guiado del cuerpo rígido, es la parte de la síntesis de mecanismos que estudia el problema de situar el acoplador de un mecanismo en diversas posiciones especificadas. Este tipo de síntesis se desarrolla en el capítulo quinto, y se inicia con dos métodos gráficos, el del polo y el de reducción de posiciones, muy útiles y sencillos, para la síntesis de guiado de la biela del cuadrilátero articulado, en tres y cuatro posiciones respectivamente. El mayor interés de esta síntesis se centra en la teoría de Burmester para el guiado de la biela en cuatro o cinco posiciones. En esta teoría se definen y hace uso de las curvas de Burmester: la "curva de los puntos circulares" o lugar geométrico de los puntos del plano móvil, que, para cuatro posiciones de éste, están sobre una circunferencia, y la "curva de centros", o lugar geométrico de los puntos del plano fijo que son centros de las anteriores circunferencias. Seguidamente se exponen los dos métodos analíticos, el de Hackmueller y el de Artobolevskii, que nos dan la ecuación de la curva de centros, así como el de la ecuación de Kaufman, que permite determinar la curva de los puntos circulares y la de centros. Finalmente, se expone un método gráfico para determinar las curvas de Burmester.

La "Optimización en síntesis", como su nombre indica, tiene por objeto buscar entre todas las posibles soluciones técnicamente válidas aquella que teniendo presente los medios disponibles y las exigencias del diseño, pueda considerarse la "mejor". Este es el problema que el autor presenta en el sexto y último capítulo de la obra.

La optimización en la síntesis de mecanismos, como consecuencia del mayor número de condiciones exigidas al mismo, al de condiciones permitidas por éste, se plantea el problema de optimizar haciendo mínima una cierta función de error. El doctor Nieto Nieto, como él indica, no pretende tratar profundamente este problema, y reseña para las personas que se sientan interesadas en él, la bibliografía especializada. Estudia la optimización de la síntesis de generación de funciones y de trayectorias en el cuadrilátero articulado, por medio de la ecuación de Freudenstein y de la técnica del error mínimo cuadrático, así como el método general de Han, que utiliza desarrollos en serie de Taylor para linealizar la función de error, para aquellos casos en que las derivaciones parciales que minimizan el error, no dan un sistema de ecuaciones lineales.

Consideramos que esta obra merece ocupar un lugar destacado en la biblioteca técnica de los especialistas en mecánica y mecanismos. (L. A.)

INTRODUCCION AL ESTADO CRISTALINO, por P. Gay. Editorial Universitaria de Barcelona, 1978. Ya en el prefacio nos damos cuenta de que el libro "Introducción al estado cristalino" va dedicado fundamentalmente a estudiantes. El autor ha sido profesor de cristalografía durante más de veinte años en Cambridge. No hay que buscar, pues, un exceso de rigor científico; el enunciado estrictamente matemático de la materia se ha recortado deliberadamente en aras de una más fácil comprensión por parte del alumno. Esta clase de libros adquieren su significación en el entorno definido de un curso de facultad o escuela universitaria y pierden un poco su importancia al situarlos en un contexto general bibliográfico. No por ello deja de ser interesante su aparición como opción de trabajo para los estudiosos de cristalografía. Y hemos de agradecer el esfuerzo del traductor para ponerlo al alcance de los lectores de lengua española. Quizá la obra llegue con retraso, no tanto por la fecha de su primera edición, sino cuanto por la estructura global de la misma.

Hay en ella un excesivo peso de los conceptos geométricos del cristal y falta mayor énfasis en el estudio del cristal real. La obra se desarrolla sobre tres ideas-tipo: la geometría cristalina, el método de estudio de la materia cristalina y el cristal real. El primer capítulo está dedicado a averiguar qué es la

materia cristalina, cómo definirla, dónde empieza, dónde termina, es decir: cuáles son los límites de la misma. Si la materia cristalina viene definida por una ordenación de sus partículas constitutivas, la cristaloquímica es la parte de la cristalografía que estudia los tipos de ordenación. Gay, a modo de ejemplo, muestra dos estructuras: la casiterita y la urea, una inorgánica y otra orgánica; define entonces los conceptos de coordinación, coordenadas atónicas y contenido de la celda. Ambos ejemplos no tienen más propósito que indicar al lector cuál va a ser el objeto real de estudio. Acaba el capítulo con una bibliografía seleccionada que creemos francamente oportuna, aun cuando deje de mencionarse las revistas de la especialidad.

El núcleo de la obra empieza con los modelos bidimensionales. Los modelos repetitivos dan lugar a la definición de dos vectores no paralelos, denominados translaciones fundamentales, que definirán la llamada "celda-unidad" o malla unidad, cuya repetición cubrirá todo el espacio. Esta malla está limitada por los vectores de menor longitud, por lo que es única. Cualquier otra celda es múltiple de la celda así definida, y recibe el nombre "celda-múltiple". En el espacio generado por la repetición infinita, en dos dimensiones, de la celda unidad, se pueden definir los llamados operadores o agentes generadores de la red, ya que los modelos pueden contener elementos repetitivos de otra clase más distantes. Los ejes de rotación, las líneas de reflexión, reflexión más translación y rotación más inversión son los operadores que pueden definirse sobre un modelo bidimensional.

Gay desarrolla las leyes de asociación de estos operadores, así como la demostración de que los ejes de rotación no pueden ser más que de orden $n + 1$, 2, 3, 4 y 6. La extensión de la simetría de dos dimensiones a la de tres nos lleva al conocimiento de los tipos de redes bidimensionales, donde, por analogía, se especificarán los diferentes tipos de celdas unidad.

Los grupos puntuales bidimensiones vienen definidos como los diferentes conjuntos de elementos de simetría, sin repetición translacional. Estamos en presencia de grupos de operaciones cuyo simbolismo viene asignado al orden del eje y a la disposición del plano: así el grupo 3*m*, tiene un eje de orden 3 y 3 planos (*m*) que contienen este eje.

Notamos a faltar un desarrollo más amplio de esta notación, denominada de Hermann-Maugin, así como de las leyes que regulan la asociación de los operadores de simetría. Finaliza el capítulo con los modelos bidimensiones espaciales y su extensión a modelos tridimensionales. Este capítulo, base de toda la cristalografía, donde se fundan y convergen todas las líneas del conocimiento cristalino, es de primordial importancia. Lamentamos que Gay exponga conceptos trascendentales de manera tan confusa, literaria, apoyándose más en una forma verbal de escritura que en un rigorismo expositivo. Es difícil calibrar, para un estudiante, cuáles son los conceptos básicos, pues se hallan mezclados en un impersonalizado conjunto de ejemplos. La omisión voluntaria de las bases matemáticas lleva a Gay a expresar las definiciones mediante palabras, con lo que no estamos de acuerdo, y menos remitiendo al lector a un libro tan conceptual como *Los elementos de la cristalografía* (Elementary crystallography) de M. J. Buerger.

El capítulo 3 expone y desarrolla la proyección estereográfica, medio de representación de los cristales y la simetría que contienen. Sobre esta fase se está en disposición para describir cada una de las clases de simetría y sistema cristalinos. La agrupación en sistemas es clásica y suscribimos el criterio de introducir un sistema trigonal en los sistemas cristalinos. Son realmente notables los estereogramas, con los que se relata la derivación de las clases cristalinas; tienen belleza y claridad. El simbolismo de los planos reticulares y de los ejes de zona requieren una previa redefinición de lo que es materia cristalina, que Gay introduce en el capítulo 5.

La expresión vectorial de la materia cristalina es un hecho impuesto al autor, "a posteriori", cuando autores tan prestigiosos como Buerger o Amorós lo introducen en el primer capítulo. Le es impuesto, porque trata de establecer números, que expresan relaciones analíticas concretas con sistemas coordenados. El libro adolece aquí de falta de actualidad; da la sensación de que el capítulo quinto ha sido añadido sobre un texto antiguo, introduciendo las nuevas ideas sobre pedagogía cristalográfica. Este apartado ocupa los cálculos de índices de plano y zona. Interés merece un largo desarrollo sobre cambios de coordenadas y sistemas de ejes alternativos.

La morfología cristalográfica es el tema de un capítulo, del cual extraemos la siguiente cita: "... o sesenta años, la difracción de rayos X y otras técnicas han sido utilizadas para la descripción de posiciones atómicas internas, e inevitablemente la importancia de la forma cristalina para el cristalógrafo moderno ha ido decreciendo...". Parece probable que Gay dudara de la eficacia de este capítulo; se trata de la descripción de algo históricamente importantísimo, pero periclitado ya en cristalografía.

Por la mitad del libro nos encontramos con un capítulo que en casi todos los tratados de cristalografía se toca de manera muy superficial debido a la enorme dificultad y aridez. Nos referimos a los tipos de redes tridimensionales. Gay los describe ayudando al texto con numerosos grabados, con lo que la exposición gana en claridad. Después de describir los operadores de simetría espacial como los ejes helicoidales, deriva el grupo *Pccn* sobre el simbolismo diagramático de International Tables of X-Ray Cristallography y lo estudia tanto en la disposición como en el contenido de los motivos estructurales. Nos gusta el capítulo por la aportación esquemática que presenta.

Con el descubrimiento de los rayos X por Roentgen en 1895 y el experimento de Von Laue se iniciaba la llamada cristalografía moderna. Tal hecho condiciona un libro de cristalografía y adquiere singular desarrollo el tema propuesto. Gay dedica tres capítulos, con un total de 84 páginas, a su exposición. Empieza con la naturaleza de los rayos X, cómo se producen, dónde se producen y resuelven: una larga digresión sobre los tubos de rayos X. En el mismo capítulo desarrolla los principios ópticos de la difracción: "de hecho, los principios ópticos de la difracción de la luz producida por una disposición regular de difusores son semejantes a los que controlan la difracción de rayos X a los cristales". Es evidente que esto es así, y resulta de interés que Gay proponga una explicación del fenómeno de la difracción a través de su análogo óptico. Podría haber incluido la exposición experimental conveniente para el lector. A partir de las figuras de difracción de Franhofer desarrolla la difracción de los rayos X en los cristales, llegando a la fundamental ley de Bragg $n\lambda = 2d \operatorname{sen} \theta$, base de toda la cristalografía de rayos X. Con las manchas de difracción se llega al cálculo de un grupo espacial, cuya significación se adquiere en el desarrollo del factor de estructura geométrico, tal como expone el autor. Este capítulo se cierra con una amplia bibliografía. Estamos convencidos de las dificultades que implica proveer un tema básico, como la difracción de rayos X, de autores significativos y apropiados para la información del lector; la selección es ajustada y correcta.

El capítulo noveno se dedica a los mé-

todos de difracción de cristal único, estáticos y dinámicos. El método de Laue se realiza con el cristal quieto y radiación blanca. En los métodos de rotación y oscilación, el cristal gira y la radiación es monocromática. Se exponen asimismo los métodos de precesión y el de Weissemberg, englobados en otros. Lamentamos que ambos, esenciales en cristalografía, reciban un trato tan discriminatorio como el dado por el autor. Un método universalmente empleado para la resolución de estructura, el difractómetro automático o de cuatro limbos, ni se menciona. Lo consideramos una grave omisión. Con correspondientes gráficas explicativas, se aporta unas fotografías de cada método que no interpreta, y sería de desear que sí lo fuesen para facilitar la tarea del lector. Los métodos de polvo se exponen con gran lujo de comentarios, tanto en lo referente a los métodos propiamente dichos, como en cuanto a la interpretación de los resultados. Diferimos en la validez de los ejemplos y sobre todo de afirmaciones tales como: "los coeficientes de dilatación se pueden obtener de manera sencilla con el método de polvo", afirmación totalmente gratuita, ya que si teóricamente es sencilla la expresión que relata el fenómeno, en la práctica se requiere gran experiencia y precisión para llegar a obtener valores significativos.

El capítulo 11 describe la relación existente entre la simetría y las propiedades físicas. Cuando un autor pretende condensar en 12 páginas un conjunto de temas, que cada uno de ellos es motivo de un libro de 500, se arriesga a dejar escrito solamente breves connotaciones de los mismos sin profundizar en ninguno. Además, no se puede hablar de propiedades físicas sin describir antes el desarrollo matemático de tensores. Por otra parte, al emplear palabras para describir un fenómeno físico se cae en el defecto de la inconcreción, tal como ocurre en este libro. Capítulo confuso, imposible de ser comprendido, si no es por un especialista impuesto en el asunto en cuestión.

Solamente el tema de óptica cristalina está desarrollado con claridad, aunque de manera escueta. Al leer la bibliografía, nos damos cuenta del valor atribuido a la óptica, en detrimento de las demás propiedades físicas, pues los textos generales constituyen solamente el 30 por ciento de los textos de consulta. Repetidas veces nos encontramos, a lo largo del libro, con temas que por su interés constituyen materia de un libro único. Este capítulo lo demuestra con creces.

Cabría preguntarnos si tiene sentido una visión totalmente superficial de un tema porque el sentido global de la materia obligue a tocarlo. Quizá sería mejor abordar un conjunto de temas, tratados en profundidad, que no el añadido de un par de capítulos a modo de addenda porque la cuestión lo requiera. Ello se hace patente al leer el capítulo dedicado a imperfecciones en los cristales reales. Está el libro descompensado, si se tiene en cuenta que tan importante es el desarrollo de los conceptos de la materia cristalina como el estudio de partes de materia cristalina concretizadas en el cristal. Se repite lo observado a lo largo del libro: brevedad, ambigüedad, abundancia de letra escrita en detrimento de lo concepual y, sobre todo, omisiones: cristal ideal y real han de tener igual número de capítulos y el balance es de siete a uno.

El libro termina con cuatro apéndices dedicados a las proyecciones y sus propiedades, cálculo sobre el estereograma, crecimiento de cristales. Son en definitiva cuatro apéndices destinados a desarrollar prácticas sencillas de cálculo: relaciones paramétricas, índices de Miller, ángulos entre zonas cristalográficas, cristalización de sustancias y morfología de las mismas. A pesar de su título genérico, son la "teoría" de las prácticas, que se complementan a lo largo del discurso, con los problemas específicos de cada lección. Estos merecen un comentario por la disparidad de dificultades que presentan: al lado de problemas de solución inmediata, hay problemas de alta dificultad resolutiva. Otras veces el autor emplea la fórmula recurrente de "hallar las deducciones que se pueda", pregunta ambigua, que se presta a soluciones heterodoxas.

La valoración global del contenido se tiene que hacer con lo que dice y lo que no dice, ya que en las primeras líneas de este comentario afirmamos que es un libro para y por los estudiantes. A su favor tiene el desarrollo profundo de la cristalografía geométrica; en contra, la falta de mayor extensión en los conceptos de cristal real. Con algunos fallos de léxico, está bien traducido y pulcramente presentado; quizá se eche de menos un comentario "avant-propos" del traductor o del adaptador; en todo caso, es un libro para añadir a la biblioteca de cristalografía. (J. A. C.)

CIRUGIA PEDIATRICA. ESQUEMAS DE TRATAMIENTO Y DIAGNÓSTICO, por M. Bettex, F. Kuffer y A. Scharli. Ed. Toray-Masson, S. A. Barcelona, 1978. Como el propio subtitulo de la portada indica, se trata de un compendio de esquemas de diagnóstico y tratamiento de los principales problemas quirúrgicos que se presentan en la práctica pediátrica. Los autores, profesores del prestigioso Universitats-Kinalerklinik Inselspital de Berna y del Chirurgischen Abteilung des Kinderspital de Lucerna, no han pretendido ofrecer un tratado de cirugía pediátrica semejante a los ya existentes en la literatura, sino editar correctamente los esquemas de consulta que, en forma de manuscrito, utilizan los estudiantes de medicina de la Universidad de Berna; así lo expone el prólogo. Esta concepción de la obra le confiere un especial atractivo. En efecto, ante un texto de conocida solvencia, resulta sumamente práctico leer en pocas líneas la definición del problema, los criterios diagnósticos a seguir y la conducta o conductas terapéuticas más aconsejables según los casos.

El libro está dividido en once apartados. Tras un capítulo de generalidades, aborda sucesivamente los procesos quirúrgicos más frecuentes de la cabeza, cara, cuello, tórax, diafragma, abdomen, tracto urogenital, columna vertebral y medula espinal, piel y partes blandas; por último, dedica un amplio capítulo a la ortopedia pediátrica.

Para cada tema se exponen en primer lugar las generalidades (definición, fisiopatología, etiología, frecuencia, etcétera) para pasar a continuación a la clínica (anamnesis, exploración, exámenes complementarios y evolución) y al tratamiento. Se evitan en general las disquisiciones y comentarios extensos, lo cual hace que en ocasiones el lector sienta la necesidad de ampliar el tema a través de otros textos. Los esquemas son sumamente didácticos y la iconografía muy cuidada. Las técnicas quirúrgicas sólo se exponen en determinados casos, pero siempre quedan claros la necésidad o no de intervención y su calendario. La traducción es correcta pese a algunos términos confusos, en especial en el capítulo primero. La edición es excelente.

En resumen, se trata de un libro que sin duda llenará un hueco importante en la literatura pediátrica española. Es muy útil y aconsejable para los estudiantes de pediatría, para los pediatras generales y para el médico de familia. En un reducido volumen puede encontrarse una respuesta rápida y concisa a los problemas de patología quirúrgica pediátrica más frecuentes, bien entendido que, una vez conocidos los criterios generales, en la mayoría de las ocasiones dicha información deberá ser completada mediante otros textos más amplios. (J. A.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

MATRIMONIO Y FERTILIDAD EN LOS PAISES DESARROLLADOS

FERTILITY DECLINE IN THE UNITED STATES AND ITS IMPLICATIONS. Charles F. Westoff en *Demographic Dynamics in America*, por Wilbur J. Cohen y Charles F. Westoff. The Free Press, 1977.

SOME SPECULATIONS ON THE FUTURE OF MARRIAGE AND FERTILITY. Charles F. Westoff en *Family Planning Perspectives*, vol. 10, n.º 2, págs. 79-83; marzo/abril, 1978.

THE PREDICTABILITY OF FERTILITY IN DEVELOPED COUNTRIES. Charles F. Westoff en *Population Bulletin of the United Nations*, n.º 11, 78. XIII. 7, en prensa.

MOLECULAS NUCLEARES

RESONANCE ENERGIES AND PARTIAL WIDTHS OF QUASIMOLECULAR STATES FORMED BY THE TWO CARBON NUCLEI. B. Imanishi en *Physics Letters*, vol. 27B, n.º 5, págs. 267-270; 5 de agosto de 1968.

QUASIMOLECULAR STRUCTURE IN ELASTIC $O^{16} + O^{17}$ SCATTERING. Werner Scheid, Walter Greiner y Richard Lemmer en *Physical Review Letters*, vol. 25, n.º 3, págs. 176-180; 20 de julio de 1970.

POSSIBLE QUASI-MOLECULAR ROTATIONAL BANDS IN THE SD-SHELL NUCLEI. A. Arima, Gertrude Scharff-Goldhaber y K. W. McVoy en *Physics Letters*, vol. 40B, n.º 1, págs. 7-10; 12 de junio de 1972.

NUCLEAR MOLECULAR STRUCTURE IN ^{12}C-^{12}C SCATTERING. H.-J. Fink, W. Scheid y W. Greiner en *Nuclear Physics*, vol. A188, págs. 259-288; junio-julio, 1972.

CENTRALES ENERGETICAS DE CELULAS DE COMBUSTIBLE

FUEL CELLS AND FUEL BATTERIES: A GUIDE TO THEIR RESEARCH AND DEVELOPMENT. Herman A. Liebhafsky y E. J. Cairns. John Wiley & Sons, 1968.

FUEL CELLS FOR PUBLIC UTILITY AND INDUSTRIAL POWER. Dirigido por Robert Noyes. Noyes Data Corporation, 1977.

LA ESTRUCTURA DE LA HEMOGLOBINA Y EL TRANSPORTE RESPIRATORIO

HEMOGLOBIN AND MYOGLOBIN IN THEIR REACTIONS WITH LIGANDS. Eraldo Antonini y Maurizio Brunori. North-Holland Publishing Company, 1971.

MAN'S HAEMOGLOBINS. H. Lehmann y R. G. Huntsman. J. P. Lippincott Company, 1974.

HUMAN HEMOGLOBINS. H. Franklin Bunn, Bernard G. Forget y Helen M. Ranney. W. B. Saunders Company, 1977.

OJOS ANIMALES DONDE LA IMAGEN SE FORMA MEDIANTE ESPEJOS

ON THE ORGANIZATION OF REFLECTING SURFACES IN SOME MARINE ANIMALS. E. J. Denton en *Philosophical Transactions of the Royal Society of London, Series B*, vol. 258, págs. 285-313; 14 de mayo de 1970.

THE PHYSICS AND BIOLOGY OF ANIMAL REFLECTORS. M. F. Land en *Progress in Biophysics and Molecular Biology*, vol. 24, págs. 75-106; 1972.

EL TEMPLO ARRASADO DE AKHENATON

THE AMARNA PERIOD AND THE END OF THE EIGHTEENTH DYNASTY. Cyril Aldred en *Cambridge Ancient History, Vol. 2, Part 2: History of the Middle East and the Aegean Region c. 1380-1000 B.C.*, dirigido por I. E. S. Edwards, C. J. Gadd, N G. L. Hammond y E. Sollberger. Cambridge at the University Press, 1975.

RECONSTRUCTING THE TEMPLES OF A HERETICAL PHARAOH. Donald B. Redford en *Archaeology*, vol. 28, n.º 1, págs. 16-22; enero, 1975.

THE EXCAVATIONS OF THE UNIVERSITY MUSEUM AKHENATEN TEMPLE PROJECT AT KARNAK. Donald B. Redford en *Expedition*, vol. 19, n.º 4, págs. 33-38; verano, 1977.

NEBULOSAS GASEOSAS

GALACTIC AND EXTRA-GALACTIC RADIO ASTRONOMY. Dirigido por Gerrit L. Verschuur y Kenneth I. Kellerman. Springer-Verlag, 1974.

A MICROWAVE INVESTIGATION OF THE TRIFID NEBULA AND ITS SURROUNDING ENVIRONMENT. Eric J. Chaisson y Robert F. Willson en *The Astrophysical Journal*, vol. 199, págs. 647-659; 1 de agosto de 1975.

GASEOUS NEBULAE AND THEIR SURROUNDING ENVIRONMENT. Eric J. Chaisson en *Frontiers of Astrophysics*. Harvard University Press, 1976.

METABOLISMO MOLECULAR DE LA TRANSDUCCION DE ENERGIA

MEMBRANE ADENOSINE TRIPHOSPHATASE OF *MICROCOCCUS LYSODEIKTICUS*. MOLECULAR PROPERTIES OF THE PURIFIED ENZYME UNSTIMULATED BY TRYPSIN. José Manuel Andreu, J. Antonio Albendea y Emilio Muñoz en *European Journal Biochemistry*, vol. 3, págs. 505-515; septiembre, 1973.

MITOCHONDRIAL ADENOSINE TRIPHOSPHATASE. Peter L. Pedersen en *Bioenergetics*, vol. 6 n.º 6, págs,. 243-275, febrero, 1975.

STRUCTURE AND FUNCTION OF CHLOROPLAST ATPASE. Nathan Nelson en *Biochimica et Biophysica Acta*. vol. 456, n.º 3/4, págs. 314-338; noviembre, 1976.

H^+-ADENOSINE TRIPHOSPHATASE AND MEMBRANE ENERGY COUPLING. I. A. Kozlov y V. P. Skulachev en *Biochimica et Biophysica Acta*, vol. 463, n.º 1, págs. 29-89; junio, 1977.

COMO FABRICAN ATP LAS CELULAS. Peter C. Hinkle y Richard E. McCarty en *Investigación y Ciencia*, n.º 20, págs. 58-75; mayo, 1978.

JUEGOS MATEMATICOS

SEMANTIC INFORMATION PROCESSING. Dirigido por Marvin Minsky. The MIT Press, 1968.

UNDERSTANDING NATURAL LANGUAGE. Terry Winograd. Academic Press, 1972.

COMPUTER MODELS OF TOUGHT AND LANGUAGE. Dirigido por Roger C. Schank y Kenneth Mark Colby, W. H. Freeman and Company, 1973.

TALLER Y LABORATORIO

PAPER MOVIE MACHINES. Budd Wentz. Troubadour Press, 1975.

MAKING VICTORIAN KINETIC TOYS. Philip y Caroline F. Sayer. Taplinger Publishing Co., 1977.

READY-TO-MAKE PHOTO & SCENE MACHINES. Budd Wentz. Troubadour Press, 1977.